



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 116605

TO: Jeffrey Parkin
Location: rem/3d39/3c18
Art Unit: 1648
March 25, 2004

Case Serial Number: 10/076421

From: P. Sheppard
Location: Remsen Building
Phone: (571) 272-2529

sheppard@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

116605

From: Parkin, Jeffrey
Sent: Wednesday, March 10, 2004 10:14 PM
To: STIC-Biotech/ChemLib
Subject: U.S. Serial No. 10/076,421

Would you please search **SEQ ID NOS.: 1-5**, from **U.S. Serial No. 10/076,421**, v. all relevant databases, including interference.

Place results on BOTH paper and electronic format.

Thanks!

JSP
AU 1648
REM 3D39
2-0908

RECEIVED
MAR 11 2004
STIC-Biotech Division
(STIC)

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 3/25/04
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 17:00:24 ; Search time 5347 Seconds

(without alignments)
10505.437 Million cell updates/sec

Title: US-10-076-421-1

Perfect score: 1296

Sequence: 1 atgagagaccctgctgscgcgcg.....agatgacctgccccctga 1296

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6340544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_ba:*
3: gb_ba:*
4: gb_ba:*
5: gb_ba:*
6: gb_ba:*
7: gb_ba:*
8: gb_ba:*
9: gb_ba:*
10: gb_ba:*
11: gb_ba:*
12: gb_ba:*
13: gb_ba:*
14: gb_ba:*
15: gb_ba:*
16: gb_ba:*
17: gb_ba:*
18: gb_ba:*
19: gb_ba:*
20: gb_ba:*
21: gb_ba:*
22: gb_ba:*
23: gb_ba:*
24: gb_ba:*
25: gb_ba:*
26: gb_ba:*
27: gb_ba:*
28: gb_ba:*
29: gb_ba:*
30: gb_ba:*
31: gb_ba:*
32: gb_ba:*
33: gb_ba:*
34: gb_ba:*
35: gb_ba:*
36: gb_ba:*
37: gb_ba:*
38: gb_ba:*
39: gb_ba:*
40: gb_ba:*
41: gb_ba:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1296	100.0	1296	6	AX512750
2	1296	100.0	1296	6	BD183647
3	1296	100.0	1296	6	BD275726
4	1296	100.0	1296	6	AR220511
5	1296	100.0	1296	6	AR255505
6	1296	100.0	1296	6	AR281075
7	1296	100.0	1296	6	AR380672
8	1296	100.0	1296	6	AR437866
9	1296	100.0	1296	6	AX365730
10	1296	100.0	1296	6	BD226055
11	1296	100.0	1296	9	HUMTUKA
12	1296	100.0	1296	11	G27040
13	1296	99.9	1296	6	A35395
14	1296	99.9	1296	6	BC013575
15	1296	99.8	1296	6	E02577
16	1296	99.8	1296	6	E02647
17	1296	99.8	1296	6	E02711
18	1296	99.8	1296	6	E06064
19	1296	99.8	1296	6	E00421
20	1296	99.8	1296	6	E01559
21	1296	99.8	1296	6	E02095
22	1296	99.8	1296	6	BD275725
23	1296	99.8	1296	6	E01560
24	1296	99.8	1296	6	AR220510
25	1296	99.8	1296	6	AR255504
26	1296	99.8	1296	6	AR281074
27	1296	99.8	1296	6	AR437865
28	1296	99.8	1296	6	AX365729
29	1296	99.8	1296	6	BD226054
30	1296	99.8	1296	9	HUMTUKA
31	1296	99.8	1296	6	A10915
32	1296	99.8	1296	6	A10916
33	1296	99.8	1296	6	A18397
34	1296	99.8	1296	6	AX402384
35	1296	99.8	1296	6	AX451989
36	1296	99.8	1296	6	E01467
37	1296	99.7	1296	9	BT007391
38	1296	99.6	1296	6	E02649
39	1296	99.6	1296	6	E00924
40	1296	99.6	1296	6	E01580
41	1296	99.6	1296	6	A11978
42	1296	99.6	1296	6	A09202
43	1296	99.6	1296	6	A21571
44	1296	99.6	1296	12	AY335599
45	1296	99.5	1340	9	HSPAVKR

ALIGNMENTS

RESULT 1
AX512750 1296 bp DNA linear PAT 03-OCT-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

Sequence 1 from Patent EP1232755...
AX512750.1 GI:23503968
Homo sapiens (human)
Homo sapiens
Bukaryota, Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Wada, M. and Wada, N.
Anti-hiv agents
Patent: EP 1232755-A 1 21-AUG-2002;

Pred. No. is the number of results predicted by chance to have a

Db 61 AGCAATGAACTTCATCAAGTTCATGCACTGTCTAATGAGAGCAATGCTG 120
 QY 121 TCCAAAGAACTTCTCCAACTTCACTGTGCACTGCCAAAGAAATTCGAGGGCAG 180
 Db 121 TCCAAAGAACTTCTCCAACTTCACTGTGCACTGCCAAAGAAATTCGAGGGCAG 180
 QY 181 CACTGTGAATAGATAGTCAAAACCTGTATGAGGGGAATGATCACTTTTACGAGGA 240
 Db 181 CACTGTGAATAGATAGTCAAAACCTGTATGAGGGGAATGATCACTTTTACGAGGA 240
 QY 241 AAGGCGAGCACTGACACCATGAGGCGGCTGTGCTGTGCACTGTGCACTGTCTT 300
 Db 241 AAGGCGAGCACTGACACCATGAGGCGGCTGTGCTGTGCACTGTGCACTGTCTT 300
 QY 301 CAGCAACGTAACATGCGCCACATCTGATGCTTTCAGCTGGGCGGGAACATTAAT 360
 Db 301 CAGCAACGTAACATGCGCCACATCTGATGCTTTCAGCTGGGCGGGAACATTAAT 360
 QY 361 TACTGAGAGAACCCAGACCAACCGGAGCGACCTGTGCTATGTGCAAGTGGGCTTAAG 420
 Db 361 TACTGAGAGAACCCAGACCAACCGGAGCGACCTGTGCTATGTGCAAGTGGGCTTAAG 420
 QY 421 CCGCTTGTCCAGAGTGCATGTGATGATCTGCGCAGATGGAAGAAAGCCCTCTCTCT 480
 Db 421 CCGCTTGTCCAGAGTGCATGTGATGATCTGCGCAGATGGAAGAAAGCCCTCTCTCT 480
 QY 481 CCGAAGAAATTAATTTTCACTGTGCGCAAAACCTCTGAGGCGGCTTTAAGATTAT 540
 Db 481 CCGAAGAAATTAATTTTCACTGTGCGCAAAACCTCTGAGGCGGCTTTAAGATTAT 540
 QY 541 GGGGGAGAAATTCACCAACATCGAGAACCGGCTGTGGGCGGCACTTACAGAGGAC 600
 Db 541 GGGGGAGAAATTCACCAACATCGAGAACCGGCTGTGGGCGGCACTTACAGAGGAC 600
 QY 601 CCGGGGGGGCTGTCTCACTGTGTGTGAGGCGGCTCTACGCCCTTGTGGGTGATC 660
 Db 601 CCGGGGGGGCTGTCTCACTGTGTGTGAGGCGGCTCTACGCCCTTGTGGGTGATC 660
 QY 661 AAGGCGACACACTGCTTCACTGTTTACCCAAAGAGAGACTACATCGCTACGCGGT 720
 Db 661 AAGGCGACACACTGCTTCACTGTTTACCCAAAGAGAGACTACATCGCTACGCGGT 720
 QY 721 CCGTCAAGGCTTAACCTCAACACGCAAGGGGAGATAGTTGAGTGGGAAACCTAATC 780
 Db 721 CCGTCAAGGCTTAACCTCAACACGCAAGGGGAGATAGTTGAGTGGGAAACCTAATC 780
 QY 781 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACCAACAGCACTTGTGCTGAGG 840
 Db 781 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACCAACAGCACTTGTGCTGAGG 840
 QY 841 ATCCGTTCCAGAGAGGGGAGGTGTGCGACCATCCCGGCTATACAGACCATCTGCTG 900
 Db 841 ATCCGTTCCAGAGAGGGGAGGTGTGCGACCATCCCGGCTATACAGACCATCTGCTG 900
 QY 901 CCTCGATGTATACAGATCCCGAGTTTGGACAAGGTGTGAGATCACTGGCTTTGAGAAA 960
 Db 901 CCTCGATGTATACAGATCCCGAGTTTGGACAAGGTGTGAGATCACTGGCTTTGAGAAA 960
 QY 961 GAGAAATTTACAGACTATCTATCCGAGACAGCTGAAAATGACTTTTGAAGCTGATT 1020
 Db 961 GAGAAATTTACAGACTATCTATCCGAGACAGCTGAAAATGACTTTTGAAGCTGATT 1020
 QY 1021 TCCCAACGGGAGGTGTGAGAGCCCACTACAGGCTGTGAGTCAACCAAAATGCTG 1080
 Db 1021 TCCCAACGGGAGGTGTGAGAGCCCACTACAGGCTGTGAGTCAACCAAAATGCTG 1080
 QY 1081 TGTGCTGTGACCCAGAGTGAAGAAAGATTCTGTGCAAGGAGACTCAAGGGGAGCCCTTC 1140
 Db 1081 TGTGCTGTGACCCAGAGTGAAGAAAGATTCTGTGCAAGGAGACTCAAGGGGAGCCCTTC 1140
 QY 1141 GTCTGTGTCCCTCCAAAGCGGCACTTGTGACTGGAATGTGAGCTGGGGCGGTGATGT 1200
 Db 1141 GTCTGTGTCCCTCCAAAGCGGCACTTGTGACTGGAATGTGAGCTGGGGCGGTGATGT 1200

QY 1201 GCCCTGAGAGACAGGCGGCGCTACAGAGAGTCTCACTTTTACCCTGGATCCGC 1260
 Db 1201 GCCCTGAGAGACAGGCGGCGCTACAGAGAGTCTCACTTTTACCCTGGATCCGC 1260
 QY 1261 AGTCACACCAAGAGAGAGATGCGCTGCGCTCTGA 1296
 Db 1261 AGTCACACCAAGAGAGAGATGCGCTGCGCTCTGA 1296
 RESULT 3
 BD275726 2294 bp DNA linear PAT 17-JUL-2003
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified.
 unclassified.
 unclassified.
 1 (bases 1 to 2294)
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 OS Homo sapien
 PN JP 2002543769-A/120
 PD 24-DEC-2002
 PE 03-APR-2000 JP 200611554
 PR 22-FEB-2000 US 09/510376,10-JAN-2000 US 09/480884, PR
 02-APR-1999 US 09/285479,30-DEC-1999 US 09/476496, PR
 17-DEC-1999 US 09/466396
 PI longlong wang, rtkn van
 CC
 FEATURES
 source
 1..2294
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 100.0%; Score 1296; DB 6; Length 2294;
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 1296; Conservative 0;

QY 1 ATGAGAGCCCTGCTGCGGCGCTTCTCTGCGTCTGTGAGCGACTCCAAAGGC 60
 Db 67 ATGAGAGCCCTGCTGCGGCGCTTCTCTGCGTCTGTGAGCGACTCCAAAGGC 126
 QY 61 AGCAATGAACTTCATCAAGTTCATGCACTGTGATGTAATGAGAGAACTGTGTG 120
 Db 127 AGCAATGAACTTCATCAAGTTCATGCACTGTGATGTAATGAGAGAACTGTGTG 186
 QY 121 TCCAAACAATTAATTCCTCCAAACATTCACCTGCACTGCGCAAAAGAAATTCGAGGGGAG 180
 Db 187 TCCAAACAATTAATTCCTCCAAACATTCACCTGCACTGCGCAAAAGAAATTCGAGGGGAG 246
 QY 181 CACTGTGAATAGATAGTCAAAACCTGTCTGAGGGGAATGTCACCTTTTACGAGGA 240
 Db 247 CACTGTGAATAGATAGTCAAAACCTGTCTGAGGGGAATGTCACCTTTTACGAGGA 306
 QY 241 AAGGCGAGCACTGACACCAATGCGGCGGCTGTGCTGTGCACTGTGCACTGTCTT 300
 Db 307 AAGGCGAGCACTGACACCAATGCGGCGGCTGTGCTGTGCACTGTGCACTGTCTT 366
 QY 301 CAGCAACGTAACATGCGCCACATCTGATGCTTTCAGCTGGGCGGGAACATTAAT 360
 Db 367 CAGCAACGTAACATGCGCCACATCTGATGCTTTCAGCTGGGCGGGAACATTAAT 426
 QY 361 TACTGAGAGAACCCAGACCAACCGGAGCGACCTGTGCTATGTGCAAGTGGGCGCTTAAG 420
 Db 427 TACTGAGAGAACCCAGACCAACCGGAGCGACCTGTGCTATGTGCAAGTGGGCGCTTAAG 486

QY 421 CCGCTTCCAAAGTGCATGTCATGTCGCGAGATGGAAGAAAGCCCTCTCTCT 480
Db 487 CCGCTTCCAAAGTGCATGTCATGTCGCGAGATGGAAGAAAGCCCTCTCTCT 546
QY 481 CCAGAGAAATTTAAATTTCAAGTGTGGCCAAAGATCTCTGAGGCCCTTTAAAGATTAT 540
Db 547 CCAGAGAAATTTAAATTTCAAGTGTGGCCAAAGATCTCTGAGGCCCTTTAAAGATTAT 606
QY 541 GGGGAGAAATTCACCAACATCGAAGAACAGCCCTGTTGGCGGCATCTACAGAGGAC 600
Db 607 GGGGAGAAATTCACCAACATCGAAGAACAGCCCTGTTGGCGGCATCTACAGAGGAC 666
QY 601 CGGGGGGGCTCTGTACCTACCTGTCGTGTGAGAGGAGCCCTCATGAGCCCTTCTGGTGTATC 660
Db 667 CGGGGGGGCTCTGTACCTACCTGTCGTGTGAGAGGAGCCCTCATGAGCCCTTCTGGTGTATC 726
QY 661 AGCGCCACACACTGCTTCAATTTGATTTACCCAAAGAGAGGAGTACATGCTTACCTGGGT 720
Db 727 AGCGCCACACACTGCTTCAATTTGATTTACCCAAAGAGAGGAGTACATGCTTACCTGGGT 786
QY 721 CGCTCAAGGCTTAACCTCCAAACAGCGAGAGGAGATGAAAGTTGAGGTGAAAACCTATATC 780
Db 787 CGCTCAAGGCTTAACCTCCAAACAGCGAGAGGAGATGAAAGTTGAGGTGAAAACCTATATC 846
QY 781 CTACACAAAGACTACAGCGCTGACACGCTTGTCTCAACCAAGACATTTGCTGTAG 840
Db 847 CTACACAAAGACTACAGCGCTGACACGCTTGTCTCAACCAAGACATTTGCTGTAG 906
QY 841 ATCCGTTCCAAAGAGAGGAGGAGGTGTGCGACAGCCATCCGAGCTATACAGACCATCTGCTG 900
Db 907 ATCCGTTCCAAAGAGAGGAGGAGGTGTGCGACAGCCATCCGAGCTATACAGACCATCTGCTG 966
QY 901 CCCTGATGATTAACGATCCCAAGTTTGGCAAGCTGTGAGTACATCTGCTTTGGAAA 960
Db 967 CCCTGATGATTAACGATCCCAAGTTTGGCAAGCTGTGAGTACATCTGCTTTGGAAA 1026
QY 961 GAGAAATTTACCGACTATCTCTATCCGAGAGCTGAAAATGACTGTGTGAAAGCTATT 1020
Db 1027 GAGAAATTTACCGACTATCTCTATCCGAGAGCTGAAAATGACTGTGTGAAAGCTATT 1086
QY 1021 TCCCAACGGGAGTGTACAGCAAGCCCACTACACGCTCTGAAATCACCAAAATGTCTG 1080
Db 1087 TCCCAACGGGAGTGTACAGCAAGCCCACTACACGCTCTGAAATCACCAAAATGTCTG 1146
QY 1081 TGTGCTGTGACCAAGTGTGAAAACAGATTTCTGCGAGGAGACTCAGGGGAGACCCCTC 1140
Db 1147 TGTGCTGTGACCAAGTGTGAAAACAGATTTCTGCGAGGAGACTCAGGGGAGACCCCTC 1206
QY 1141 GTCTGTTCCCTCCAAAGGCGCGATGACTTTGACTGTGAATTTGTAGCTGGGCGCTGTAGT 1200
Db 1207 GTCTGTTCCCTCCAAAGGCGCGATGACTTTGACTGTGAATTTGTAGCTGGGCGCTGTAGT 1266
QY 1201 GCCCTGAAGACAAAGCCAGGCGTCTACACAGAGAGTCTCACTCTTAACTCCGTGATCCGC 1260
Db 1267 GCCCTGAAGACAAAGCCAGGCGTCTACACAGAGAGTCTCACTCTTAACTCCGTGATCCGC 1326
QY 1261 AGTCACACCAAGAGAGAGTGTGCGCTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGAGTGTGCGCTGGCCCTCTGA 1362

RESULT 4
AR220511
LOCUS AR220511 2294 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 123 from patent US 6426072.
ACCESSION AR220511
VERSION AR220511.1 GI:23327292
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2294)

AUTHORS Wang, T., Fan, L., Kalos, M. D., Bangur, C. S., Hoeken, N. A., Fanger, G. R.,
Li, S. X., Wang, A., Skelky, Y. A. W., Henderson, R. A. and McNeill, P. D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6426072-A 123 30-JUL-2002;
FEATURES Location/Qualifiers
source l . 2294
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1296; DB 6; Length 2294;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGTGCGCGCCCTGCTTCTGCTGCTGTGTGTGAGGAGCTCCAAAGGC 60
Db 67 ATGAGAGCCCTGTGCGCGCCCTGCTTCTGCTGCTGTGTGTGAGGAGCTCCAAAGGC 126
QY 61 AGCAATGAATTCATCAAGTTCATCAAGTGTGAACTGTGCTTAAATGAGAAACAATGTGTG 120
Db 127 AGCAATGAATTCATCAAGTTCATCAAGTGTGAACTGTGCTTAAATGAGAAACAATGTGTG 186
QY 121 TCCAAAGATCTTCTCCAAACATTCAGTGTGCAACTGCCAAAGAAATTTGGAGGCGAG 180
Db 187 TCCAAAGATCTTCTCCAAACATTCAGTGTGCAACTGCCAAAGAAATTTGGAGGCGAG 246
QY 181 CACTGTGAATTAATGATTAATTCAAAACCTGCTATAGAGGAAATGTCACTTTACCGAGA 240
Db 247 CACTGTGAATTAATGATTAATTCAAAACCTGCTATAGAGGAAATGTCACTTTACCGAGA 306
QY 241 AAGGCACACATGACACCATGAGCGCGCCCTGCGCTGGAGACTGTGCACTGTCTT 300
Db 307 AAGGCACACATGACACCATGAGCGCGCCCTGCGCTGGAGACTGTGCACTGTCTT 366
QY 301 CAGCAAAAGTACATGCCCAAGATCTGATGCTCTTCAAGCTGGGCTTGGGAAACATTAAT 360
Db 367 CAGCAAAAGTACATGCCCAAGATCTGATGCTCTTCAAGCTGGGCTTGGGAAACATTAAT 426
QY 361 TACTGACAGAAACCAAGAACCCGAGGCGACCCCTGGTGTATGTGACAGGTGGGCTTAAAG 420
Db 427 TACTGACAGAAACCAAGAACCCGAGGCGACCCCTGGTGTATGTGACAGGTGGGCTTAAAG 486
QY 421 CCGCTTGTCCAAAGTGCATGTGATGACTGTGCGCAGATGAAAAGAGCCCTCTCTCT 480
Db 487 CCGCTTGTCCAAAGTGCATGTGATGACTGTGCGCAGATGAAAAGAGCCCTCTCTCT 546
QY 481 CCAGAGAAATTTAAATTTCAAGTGTGGCCAAAGATCTCTGAGGCCCTGTTAAAGATTAT 540
Db 547 CCAGAGAAATTTAAATTTCAAGTGTGGCCAAAGATCTCTGAGGCCCTGTTAAAGATTAT 606
QY 541 GGGGAGAAATTCACCAACATCGAAGAACAGCCCTGTTGGCGGCATCTACAGAGGAC 600
Db 607 GGGGAGAAATTCACCAACATCGAAGAACAGCCCTGTTGGCGGCATCTACAGAGGAC 666
QY 601 CGGGGGGGCTCTGTACCTACCTGTCGTGTGAGAGGAGCTTATAGCCCTTGTGGTGTATC 660
Db 667 CGGGGGGGCTCTGTACCTACCTGTCGTGTGAGAGGAGCTTATAGCCCTTGTGGTGTATC 726
QY 661 AGCGCCACACACTGCTTCAATTTGATTTACCCAAAGAGAGGAGTACATGCTTACCTGGGT 720
Db 727 AGCGCCACACACTGCTTCAATTTGATTTACCCAAAGAGAGGAGTACATGCTTACCTGGGT 786
QY 721 CGCTCAAGGCTTAACCTCCAAACAGCGAGAGGAGATGAAAGTTGAGGTGAAAACCTATATC 780
Db 787 CGCTCAAGGCTTAACCTCCAAACAGCGAGAGGAGATGAAAGTTGAGGTGAAAACCTATATC 846
QY 781 CTACACAAAGACTACAGCGCTGACACGCTTGTCTCAACCAAGACATTTGCTGTAG 840
Db 847 CTACACAAAGACTACAGCGCTGACACGCTTGTCTCAACCAAGACATTTGCTGTAG 906
QY 841 ATCCGTTCCAAAGAGGAGGAGGTGTGCGACAGCCATCCGAGCTATACAGACCATCTGCTG 900

Db 907 ATCCGATTCAGAGGAGGAGAGGTGTGCGAGCATCCCGAGCTATACAGACCATCTGTGCTG 966
Qy 901 CCTCGATGATTAACGATCCCGAGTTTGGACAGAGCTGTAGATCACTGGCTTTGGAAAA 960
Db 967 CCTCATGATATACATCCCGAGTTTGGACAGAGCTGTAGATCACTGGCTTTGGAAAA 1026
Qy 961 GAGAAATTTACAGCATATCTCTATCCGAGAGCAGCTGAAAAATGACTGTGTGAAGCTGATT 1020
Db 1027 GAGAAATTTACAGCATATCTCTATCCGAGAGCAGCTGAAAAATGACTGTGTGAAGCTGATT 1086
Qy 1021 TCCACCGGAGAGTGTAGAGAGCCCACTACTACGGCTGTGAAGTCAACACAAAATGCTG 1080
Db 1087 TCCACCGGAGAGTGTAGAGAGCCCACTACTACGGCTGTGAAGTCAACACAAAATGCTG 1146
Qy 1081 TGTGCTGTGACCCCAAGTGTGAAGAAACAGATTCTTCCAGAGGAGACTCAGAGGAGACCCCTC 1140
Db 1147 TGTGCTGTGACCCCAAGTGTGAAGAAACAGATTCTTCCAGAGGAGACTCAGAGGAGACCCCTC 1206
Qy 1141 GTCTGTTCCCTCCCAAGCGCGCATGACTTTGACTGTGAATTGTGAGCTGGGGCGGTGAGT 1200
Db 1207 GTCTGTTCCCTCCCAAGCGCGCATGACTTTGACTGTGAATTGTGAGCTGGGGCGGTGAGT 1266
Qy 1201 GCGCTGAGAGACAGAGCGAGCGCTCTACAGAGAGTCTCACTTACCCTGAGATCCGC 1260
Db 1267 GCGCTGAGAGACAGAGCGAGCGCTCTACAGAGAGTCTCACTTACCCTGAGATCCGC 1326
Qy 1261 AGTCACACCAAGAGAGAAATGGCGCTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGAAATGGCGCTGGCCCTCTGA 1362

RESULT 5
AR255505
LOCUS AR255505 2294 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 123 from patent US 6482597.
ACCESSION AR255505
VERSION AR255505.1 GI:27304554
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2294)
AUTHORS Wang, T., Hosken, N.A., Kalos, M.D., Fanger, G.R. and Fan, L.
TITLE Compounds and methods for therapy and diagnosis of lung cancer
JOURNAL Patent: US 6482597-A 123 19-NOV-2002;
FEATURES
Location/Qualifiers
1..2294
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1296; DB 6; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGAGCCCTGTGCGCGCGCTGCTTCTGTGCTCTGTGCTGTGCGCATCTCCAAAGGC 60
Db 67 ATGAGAGCCCTGTGCGCGCGCTGCTTCTGTGCTCTGTGCTGTGCGCATCTCCAAAGGC 126
Qy 61 AGCAATGAACTTCATCAAGTTCATGAACTGTGAACTGTCTAAATGAGAGAACTATGTT 120
Db 127 AGCAATGAACTTCATCAAGTTCATGAACTGTGAACTGTCTAAATGAGAGAACTATGTT 186
Qy 121 TCCAAACAGTACTTCTCCAAATTCATCTGTGTGCACTGTGCCAAAGAAATTCGAGGGCAG 180
Db 187 TCCAAACAGTACTTCTCCAAATTCATCTGTGTGCACTGTGCCAAAGAAATTCGAGGGCAG 246
Qy 181 CACTGTGAATATGATATGTCAAAAACCTGTATGAGGGGAATGTCACTTTTACCGAGGA 240
Db 247 CACTGTGAATATGATATGTCAAAAACCTGTATGAGGGGAATGTCACTTTTACCGAGGA 306
Qy 241 AAGGCGAGACTGACACCAATGGGCGGCGCTGTGCTGAGAACTGTGCTCTT 300

Db 307 AAGGCGAGACTGACACCAATGGGCGGCGCTGTGCTGAGAACTGTGCACTGTCTT 366
Qy 301 CAGCAACGTAACCATGCGCCACAGATCTGATGCTCTTCACTGTGGGCTGTGGAAAACTAAT 360
Db 367 CAGCAACGTAACCATGCGCCACAGATCTGATGCTCTTCACTGTGGGCTGTGGAAAACTAAT 426
Qy 361 TACTGAGGAAACCCAGACAAACCGGAGGCGACCTGTGGTCTATGTGACAGGTGGGCTTAAG 420
Db 427 TACTGAGGAAACCCAGACAAACCGGAGGCGACCTGTGGTCTATGTGACAGGTGGGCTTAAG 486
Qy 421 CCGCTTGTCCAGAGTGCATGTGTGATGACTGTGCGAGATGAGAAAAAGCCCTCTCTCT 480
Db 487 CCGCTTGTCCAGAGTGCATGTGTGATGACTGTGCGAGATGAGAAAAAGCCCTCTCTCT 546
Qy 481 CCAAGAAATTAATTAATTAATGATGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGTTATT 540
Db 547 CCAAGAAATTAATTAATTAATGATGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGTTATT 606
Qy 541 GGGGGGAAATTCACCAACATCGAAGACAGCCCTGTGTTGGCGCATCTACAGAGGAC 600
Db 607 GGGGGGAAATTCACCAACATCGAAGACAGCCCTGTGTTGGCGCATCTACAGAGGAC 666
Qy 601 CCGGGGGGCTGTGTCACTACGTGTGTGAGAGGAGCGCTCATAGCCCTTGTGTGTATC 660
Db 667 CCGGGGGGCTGTGTCACTACGTGTGTGAGAGGAGCGCTCATAGCCCTTGTGTGTATC 726
Qy 661 AGGCGCAACACTGTCTTCAATGATTAACCAAGAGAGAGACTACATCGTCTACCTGGGT 720
Db 727 AGGCGCAACACTGTCTTCAATGATTAACCAAGAGAGAGACTACATCGTCTACCTGGGT 786
Qy 721 CCGTCAAGGCTTAATCCCAACCGCAAGGGGAGATGAATTTGAGTGAAGAAACCTAATC 780
Db 787 CCGTCAAGGCTTAATCCCAACCGCAAGGGGAGATGAATTTGAGTGAAGAAACCTAATC 846
Qy 781 CTACACAAAGACTACAGCGCTGACAGCTTGTCTCACACAGACTTGTGCTGTGAAG 840
Db 847 CTACACAAAGACTACAGCGCTGACAGCTTGTCTCACACAGACTTGTGCTGTGAAG 906
Qy 841 ATCCGTTCCAGAGAGGAGGTGTGTGCGAGCAATCCCGGACTATACAGACATCTGCTG 900
Db 907 ATCCGTTCCAGAGAGGAGGTGTGTGCGAGCAATCCCGGACTATACAGACATCTGCTG 966
Qy 901 CCTCGATGATTAACGATCCCGAGTTTGGACAGAGCTGTAGATCACTGGCTTTGGAAAA 960
Db 967 CCTCGATGATTAACGATCCCGAGTTTGGACAGAGCTGTAGATCACTGGCTTTGGAAAA 1026
Qy 961 GAGAAATTTACAGCATATCTCTATCCGAGAGCAGCTGAAAAATGACTGTGTGAAGCTGATT 1020
Db 1027 GAGAAATTTACAGCATATCTCTATCCGAGAGCAGCTGAAAAATGACTGTGTGAAGCTGATT 1086
Qy 1021 TCCACCGGAGAGTGTAGAGAGCCCACTACTACGGCTGTGAAGTCAACACAAAATGCTG 1080
Db 1087 TCCACCGGAGAGTGTAGAGAGCCCACTACTACGGCTGTGAAGTCAACACAAAATGCTG 1146
Qy 1081 TGTGCTGTGACCCCAAGTGTGAAGAAACAGATTCTTCCAGAGGAGACTCAGAGGAGACCCCTC 1140
Db 1147 TGTGCTGTGACCCCAAGTGTGAAGAAACAGATTCTTCCAGAGGAGACTCAGAGGAGACCCCTC 1206
Qy 1141 GTCTGTTCCCTCCCAAGCGCGCATGACTTTGACTGTGAATTGTGAGCTGGGGCGGTGAGT 1200
Db 1207 GTCTGTTCCCTCCCAAGCGCGCATGACTTTGACTGTGAATTGTGAGCTGGGGCGGTGAGT 1266
Qy 1201 GCGCTGAGAGACAGAGCGGCTCTACAGAGAGTCTCACTTTCATCTGTGATCCGC 1260
Db 1267 GCGCTGAGAGACAGAGCGGCTCTACAGAGAGTCTCACTTTCATCTGTGATCCGC 1326
Qy 1261 AGTCACACCAAGAGAGAAATGGCGCTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGAAATGGCGCTGGCCCTCTGA 1362

RESULT 6
AR281075

```

LOCUS       AR281075                2294 bp    DNA             linear      PAT 10-APR-2003
DEFINITION   Sequence 123 from patent US 6518256.
ACCESSION    AR281075
VERSION      AR281075.1 GI:29716552
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 2294)
AUTHORS     Wang, T., Fan, L., Kalos, M.D., Bangur, C.S., Hosken, N.A. and
            Fanger, G.R.
TITLE        Compounds and methods for therapy and diagnosis of lung cancer
JOURNAL     Patent: US 6518256-A 123 11-FEB-2003;
FEATURES
SOURCE      1..2294
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 1296; DB 6; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGAGAGCCCTGCTGGCGCGCTGCTTCTCTGCGTCTGCTGCTGAGCGACTCCAAAGGC 60
DB      67 ATGAGAGCCCTGCTGGCGCGCTGCTTCTCTGCGTCTGCTGAGCGACTCCAAAGGC 126

QY      61 AGCAATGAACTTCAATCAAGTTCCATCGAATCTGTAAGTCTTAAATGAGAGAACTGTGTG 120
DB      127 AGCAATGAACTTCAATCAAGTTCCATCGAATCTGTAAGTCTTAAATGAGAGAACTGTGTG 186

QY      121 TCCAAACAAGTACTTCTCCAAATTCATCATGTGTCAGTCCGCAAGAAATTCGAGGGCAG 180
DB      187 TCCAAACAAGTACTTCTCCAAATTCATCATGTGTCAGTCCGCAAGAAATTCGAGGGCAG 246

QY      181 CACTGTGAAATAGATAAGTCAAAAACTGCTATGAGGGGAATGATCACTTTTACCGANGA 240
DB      247 CACTGTGAAATAGATAAGTCAAAAACTGCTATGAGGGGAATGATCACTTTTACCGANGA 306

QY      241 AAGGCCAGCATGTCACATCATGAGGGCGGCGCTGCTGCTGCGCCCTGCGAACTCTGCACTGTCTT 300
DB      307 AAGGCCAGCATGTCACATCATGAGGGCGGCGCTGCTGCTGCGAACTCTGCACTGTCTT 366

QY      301 CAGCAAACTGATACATGCGCCACAGATCTGATCTTTCAGCTGAGGCTTGGGGAAACATTAAT 360
DB      367 CAGCAAACTGATACATGCGCCACAGATCTGATCTTTCAGCTGAGGCTTGGGGAAACATTAAT 426

QY      361 TACTGCAAGAACCCAGACAAACCGAGGCCGACCTTGCTGCTATGTGCAAGTGGGCTTAAAG 420
DB      427 TACTGCAAGAACCCAGACAAACCGAGGCCGACCTTGCTGCTATGTGCAAGTGGGCTTAAAG 486

QY      421 CCGCTTGTCCAGAGTGCATGTGTCATGATCTGCGAGATGGAAGAAAGCCCTCTCTCTCT 480
DB      487 CCGCTTGTCCAGAGTGCATGTGTCATGATCTGCGAGATGGAAGAAAGCCCTCTCTCTCTCT 546

QY      481 CCGAAGAAATTAATTTTCAAGTGTGCGCAAAAGACTCTGAGGCCCGCTTAAAGATTATT 540
DB      547 CCGAAGAAATTAATTTTCAAGTGTGCGCAAAAGACTCTGAGGCCCGCTTAAAGATTATT 606

QY      541 GGGGGGAATTTACCAACCATCGAGAACCGACCTTGCTTGGCGCATCTTACAGAGGCGAC 600
DB      607 GGGGGGAATTTACCAACCATCGAGAACCGACCTTGCTTGGCGCATCTTACAGAGGCGAC 666

QY      601 CCGGGGGGCTCTGTCACTACGTGTGAGGAGCAGCTCATGACCTTGTGCTGGGATTC 660
DB      667 CCGGGGGGCTCTGTCACTACGTGTGAGGAGCAGCTCATGACCTTGTGCTGGGATTC 726

QY      661 AGGCGCACACACTGCTTCATTGATTAACCAAGAGAGAGACTACATCTCTTACCTGGGT 720
DB      727 AGGCGCACACACTGCTTCATTGATTAACCAAGAGAGAGACTACATCTCTTACCTGGGT 786

QY      721 CGCTCAAGGCTTAATCCCAACACGCAAGGGGAGATGAAGTTTGAAGTGAAGAACTTAATC 780

```

```

DB      787 CGCTCAAGGCTTAATCCCAACACGCAAGGGGAGATGAAGTTTGAAGTGAAGAACTTAATC 846
QY      781 CTACACAAGAGACTACAGCGCTGACAGCTTGTGTCACACAAGACATTTGCCCTGTGAAG 840
DB      847 CTACACAAGAGACTACAGCGCTGACAGCTTGTGTCACACAAGACATTTGCCCTGTGAAG 906
QY      841 ATCCGTTCCAGAGAGGCGAGGTGTGCGCAGCCATCCCGACTATACAGACATCTGCCTG 900
DB      907 ATCCGTTCCAGAGAGGCGAGGTGTGCGCAGCCATCCCGACTATACAGACATCTGCCTG 966
QY      901 CCTCGATGTATTAACGATATCCCAAGTTTGGCAACAGCTGTGAGATCACTGCTTTGGAAA 960
DB      967 CCTCGATGTATTAACGATATCCCAAGTTTGGCAACAGCTGTGAGATCACTGCTTTGGAAA 1026
QY      961 GGAATTTCAACGACTATCTATCCGAGCAGCTGAAAATATGACTGTTGTGAAGCTGATT 1020
DB      1027 GGAATTTCAACGACTATCTATCCGAGCAGCTGAAAATATGACTGTTGTGAAGCTGATT 1086
QY      1021 TCCACCGGAGAGTGTACAGACGCCCACTATACGAGCTCTGAAGTACCAACCAAAATGCTG 1080
DB      1087 TCCACCGGAGAGTGTACAGACGCCCACTATACGAGCTCTGAAGTACCAACCAAAATGCTG 1146
QY      1081 TGTGCTGTGACCCCAAGTGAAGAAACAGATTCTGCGCAGGGAGACTCAAGGGGACCCCTC 1140
DB      1147 TGTGCTGTGACCCCAAGTGAAGAAACAGATTCTGCGCAGGGAGACTCAAGGGGACCCCTC 1206
QY      1141 GTCTGTTCCTTCCAAAGCGCGCATGACTTGTGAATTTGTGAGCTGGGGCGTGGATGT 1200
DB      1207 GTCTGTTCCTTCCAAAGCGCGCATGACTTGTGAATTTGTGAGCTGGGGCGTGGATGT 1266
QY      1201 GGCCTGAAGAGAACACCGCAGGCGTCTTACACGAGAGTCTCAACTTTTACCTGGATCCGC 1260
DB      1267 GGCCTGAAGAGAACACCGCAGGCGTCTTACACGAGAGTCTCAACTTTTACCTGGATCCGC 1326
QY      1261 AGTCACACCAAGAGAGAAATGAGCTGAGCCCTCTGA 1296
DB      1327 AGTCACACCAAGAGAGAAATGAGCTGAGCCCTCTGA 1362

RESULT 7
AR380672                2294 bp    DNA             linear      PAT 18-DEC-2003
LOCUS       AR380672
DEFINITION   Sequence 1217 from patent US 6607879.
ACCESSION    AR380672
VERSION      AR380672.1 GI:40089306
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 2294)
AUTHORS     Cocks, B.G., Stuart, S.G. and Seilhamer, J.J.
TITLE        Compositions for the detection of blood cell and immunological
            response gene expression
JOURNAL     Patent: US 6607879-A 1217 19-AUG-2003;
FEATURES
SOURCE      1..2294
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 1296; DB 6; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGAGAGCCCTGCTGGCGCGCTGCTTCTCTGCGTCTGCTGCTGAGCGACTCCAAAGGC 60
DB      67 ATGAGAGCCCTGCTGGCGCGCTGCTTCTCTGCGTCTGCTGAGCGACTCCAAAGGC 126

QY      61 AGCAATGAACTTCAATCAAGTTCCATCGAATCTGTAAGTCTTAAATGAGAGAACTGTGTG 120
DB      127 AGCAATGAACTTCAATCAAGTTCCATCGAATCTGTAAGTCTTAAATGAGAGAACTGTGTG 186

QY      121 TCCAAACAAGTACTTCTCCAAATTCATCATGTGTCAGTCCGCAAGAAATTCGAGGGCAG 180

```

Db 187 TCACACAGTACTTCTCCAACTTCACTGTGCACTGCCCAAGAAATTGGAGGCGAG 246
Qy 181 CACTGTGAATATAGTAAATGCAAAACCTGCTATGAGGGAGATGCTCATCTTTTACCGAGGA 240
Db 247 CACTGTGAATATAGTAAATGCAAAACCTGCTATGAGGGAGATGCTCATCTTTTACCGAGGA 306
Qy 241 AAGGCGAGCACTGACACGATGAGGCGCGGCTGCTGCTGCAATCTGTGCACTGTCTCT 300
Db 307 AAGGCGAGCACTGACACGATGAGGCGCGGCTGCTGCTGCAATCTGTGCACTGTCTCT 366
Qy 301 CAGCAACGTAACCATGCGCCACAGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAAT 360
Db 367 CAGCAACGTAACCATGCGCCACAGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAAT 426
Qy 361 TACTGACAGAACCCAGCAACCCGAGGCGACCTGTGTATGTGACAGTGGGCTTAAG 420
Db 427 TACTGACAGAACCCAGCAACCCGAGGCGACCTGTGTATGTGACAGTGGGCTTAAG 486
Qy 421 CGGCTGTCCAAAGTGCATGTGCACTGAGTGGGCAATGGAAAAACCCCTCTCTCTCT 480
Db 487 CGGCTGTCCAAAGTGCATGTGCACTGAGTGGGCAATGGAAAAACCCCTCTCTCTCT 546
Qy 481 CCAGAAAGAAATTAATTTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 540
Db 547 CCAGAAAGAAATTAATTTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 606
Qy 541 GGGGAGAAATTCACCAACCATGAGAACCAAGCCCTGTGTTGGCCCATCTACAGAGGCGAC 600
Db 607 GGGGAGAAATTCACCAACCATGAGAACCAAGCCCTGTGTTGGCCCATCTACAGAGGCGAC 666
Qy 601 CGGGGGGGCTCTGTCACTACCTGTGTGAGGAGGCAAGCTCATCAGCCCTTCTGGGTATC 660
Db 667 CGGGGGGGCTCTGTGTCACTACCTGTGTGAGGAGGCAAGCTCATCAGCCCTTCTGGGTATC 726
Qy 661 AAGGCGACACACTGCTTCATTTGATTAATCCCAAGAGAGGAGACTACATCGTCTACCTGGGT 720
Db 727 AAGGCGACACACTGCTTCATTTGATTAATCCCAAGAGAGGAGACTACATCGTCTACCTGGGT 786
Qy 721 CGGTCAGGCTTAATCTCCAAACGCAAGGAGAGATGAAGTTTGAAGTGAAGAACTTAATC 780
Db 787 CGGTCAGGCTTAATCTCCAAACGCAAGGAGAGATGAAGTTTGAAGTGAAGAACTTAATC 846
Qy 781 CTAACAAGAGACTACAGGCGTGAACGCTTGTCTACCAACGACATTTGCTTGTCTGAAG 840
Db 847 CTAACAAGAGACTACAGGCGTGAACGCTTGTCTACCAACGACATTTGCTTGTCTGAAG 906
Qy 841 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGGCAATCCCGGACTATPACAGACCATCTGCTG 900
Db 907 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGGCAATCCCGGACTATPACAGACCATCTGCTG 966
Qy 901 CCCTGATGTATACGATCCCAAGTTTGGCAAGAGCTGTGAGATCACTGGCTTTTGAAGAA 960
Db 967 CCCTGATGTATACGATCCCAAGTTTGGCAAGAGCTGTGAGATCACTGGCTTTTGAAGAA 1026
Qy 961 GAGAAATTTCAACGACTATCTATCCGAGACAGCTGAATGAAGCTGTTGTAAGCTGAT 1020
Db 1027 GAGAAATTTCAACGACTATCTATCCGAGACAGCTGAATGAAGCTGTTGTAAGCTGAT 1086
Qy 1021 TCCCAACGAGGAGTGTGAGGAGCCCACTACTACGGCTGTGAATGACCAACCAAAAGCTG 1080
Db 1087 TCCCAACGAGGAGTGTGAGGAGCCCACTACTACGGCTGTGAATGACCAACCAAAAGCTG 1146
Qy 1081 TGTGCTGTGACCCACAGTGAAGAAAGATTTCTGTGCGAGGAGACTCAAGGGGAGACCCCTC 1140
Db 1147 TGTGCTGTGACCCACAGTGAAGAAAGATTTCTGTGCGAGGAGACTCAAGGGGAGACCCCTC 1206
Qy 1141 GTCTGTTCCTTCCAAAGGCGGAGTGAATTTGAGCTGTGAGTGGGCTGTGATGT 1200
Db 1207 GTCTGTTCCTTCCAAAGGCGGAGTGAATTTGAGCTGTGAGTGGGCTGTGATGT 1266
Qy 1201 GGCCTAAAGAACAGGCGGCTCTACAGAGAGTCTCACACTTTACCTGGATCCGCG 1260

Db 1267 GCCCTGAGAGCAACGAGGCGTCTACAGAGAGTCTCACACTTTACCTGGATCCGCG 1326
Qy 1261 AGTCACACCAAGAGAGAAATGGCTTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGAAATGGCTTGGCCCTCTGA 1362
RESULT 8
AR437866
LOCUS AR437866 2294 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 123 from patent US 6660838.
ACCESSION AR437866
VERSION AR437866.1 GI:40203082
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2294)
AUTHORS Wang, T.
TITILE Compounds and methods for therapy and diagnosis of lung cancer
JOURNAL Patent: US 6660838-A 123 09-DEC-2003;
FEATURES Location/Qualifiers
source 1..2294
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1296; DB 6; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGAGCCCTGCTGTGCGCGGCTCTTCTCTGTGCTCTGTGCTGTGAGCCATCTCAAGGC 60
Db 67 ATGAGAGCCCTGCTGTGCGCGGCTCTTCTCTGTGCTCTGTGCTGTGAGCCATCTCAAGGC 126
Qy 61 AGCAATGAACTTCAATCAAGTTCATGCAACTGCACTGTGCTTAATGAGAGCAATGCTGTG 120
Db 127 AGCAATGAACTTCAATCAAGTTCATGCAACTGCACTGTGCTTAATGAGAGCAATGCTGTG 186
Qy 121 TCCAAAGTACTTCTCCAAACATTCACGTGTGCACTGCCCAAGAAATTTGGAGGCGAG 180
Db 187 TCCAAAGTACTTCTCCAAACATTCACGTGTGCACTGCCCAAGAAATTTGGAGGCGAG 246
Qy 181 CACTGTGAATATAGTAAATGCAAAACCTGCTATGAGGGAGATGCTCACTTTTACCGAGGA 240
Db 247 CACTGTGAATATAGTAAATGCAAAACCTGCTATGAGGGAGATGCTCACTTTTACCGAGGA 306
Qy 241 AAGGCGAGCACTGACACGATGAGGCGCGGCTGCTGCTGCGAATCTGTGCACTGTCTCT 300
Db 307 AAGGCGAGCACTGACACGATGAGGCGCGGCTGCTGCTGCGAATCTGTGCACTGTCTCT 366
Qy 301 CAGCAACGTAACCATGCGCCACAGATCTGATGCTTTCAGCTGGGCTGGGAAACATAAT 360
Db 367 CAGCAACGTAACCATGCGCCACAGATCTGATGCTTTCAGCTGGGCTGGGAAACATAAT 426
Qy 361 TACTGACAGAACCCAGCAACCCGAGGCGACCTGTGTATGTGACAGTGGGCTTAAG 420
Db 427 TACTGACAGAACCCAGCAACCCGAGGCGACCTGTGTATGTGACAGTGGGCTTAAG 486
Qy 421 CGGCTGTCCAAAGTGCATGTGCACTGAGTGGGCAATGGAAAAACCCCTCTCTCTCT 480
Db 487 CGGCTGTCCAAAGTGCATGTGCACTGAGTGGGCAATGGAAAAACCCCTCTCTCTCT 546
Qy 481 CCAGAAAGAAATTAATTTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 540
Db 547 CCAGAAAGAAATTAATTTTCACTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAT 606
Qy 541 GGGGAGAAATTCACCAACCATGAGAACCAAGCCCTGTGTTGGCCCATCTACAGAGGCGAC 600
Db 607 GGGGAGAAATTCACCAACCATGAGAACCAAGCCCTGTGTTGGCCCATCTACAGAGGCGAC 666
Qy 601 CGGGGGGGCTCTGTCACTACGTTGTGAGGAGGCAAGCTCATCAGCCCTTGTGGGTATC 660

Db	667	CGGGGGGGCTGTGTACCTACGTGTGTGAGGACGACCTTATCAGCCCTTGTGGTGATC	726		
Qy	661	AGGCCACACACTGCTTCTATTATTATCCCAAGAGAGGACTATCATGCTTACCTGGT	720		
Db	727	AGGGCCACACACTGTCTTCTATTATCCCAAGAGAGGAGACTATCATGCTTACCTGGT	786		
Qy	721	CGGTCAAGCTTAACTCCCAACGCAAGAGAGGAGTGAATTTGAGTGGAAACCTATATC	780		
Db	787	CGCTCAAGCTTAACTCCCAACGCAAGAGAGGAGTGAATTTGAGTGGAAACCTATATC	846		
Qy	781	CTACACAAGACTTACAGCGCTGACACGCTTGTCTACCAACAAGCATTTGCTTGTGAAG	840		
Db	847	CTACACAAGACTTACAGCGCTGACACGCTTGTCTACCAACAAGCATTTGCTTGTGAAG	906		
Qy	841	ATCCGTTCCAAAGAGAGGAGGTGTGGAGAGCCATCCGAGACTTATACAGACCATTTGCTG	900		
Db	907	ATCCGTTCCAAAGAGAGGAGGTGTGGAGAGCCATCCGAGACTTATACAGACCATTTGCTG	966		
Qy	901	CCCTGATGTATATACGATCCCAAGTTTGGACAAGCTGTGATCACTGGCTTTGAGAAA	960		
Db	967	CCCTGATGTATATATAGATCCCAAGTTTGGACAAGCTGTGATCACTGGCTTTGAGAAA	1026		
Qy	961	GAGAAATTTCAACGACTATATCTTATCCGAGACGCTGAATAATGACTGTTGTGAAGCTGAT	1020		
Db	1027	GAGAAATTTCAACGACTATATCTTATCCGAGACGCTGAATAATGACTGTTGTGAAGCTGAT	1086		
Qy	1021	TCCCAACGGGAGGTGTACAGAGCCCACTACTACGCTGTGAAGTCAACCAAAATGCTG	1080		
Db	1087	TCCCAACGGGAGGTGTACAGAGCCCACTACTACGCTGTGAAGTCAACCAAAATGCTG	1146		
Qy	1081	TGTGTGTGTGACCCCAAGTGTGAAGATTCCTGCGAGAGGAGACTCAGGGGAGACCCCTC	1140		
Db	1147	TGTGTGTGTGACCCCAAGTGTGAAGATTCCTGCGAGAGGAGACTCAGGGGAGACCCCTC	1206		
Qy	1141	GTCCTTTCCCTCCAAAGGCGCGATGCTTGAATGTGAAGCTGGGGCCCTGTGATGT	1200		
Db	1207	GTCCTTTCCCTCCAAAGGCGCGATGCTTGAATGTGAAGCTGGGGCCCTGTGATGT	1266		
Qy	1201	GCCCTGAAGACCAAGCCGAGGTGTACACGAGTGTCAACTTTCATCCCTGATCCGC	1260		
Db	1267	GCCCTGAAGACCAAGCCGAGGTGTACACGAGTGTCAACTTTCATCCCTGATCCGC	1326		
Qy	1261	AGTCACACCAAGAGAGATGGCTGGCCCTCTGA	1296		
Db	1327	AGTCACACCAAGAGAGATGGCTGGCCCTCTGA	1362		
RESULT 9	AX365730	2294 bp	DNA	linear	PAT 15-FEB-2002
LOCUS	AX365730	Sequence 123 from Patent WO0200174.			
DEFINITION	AX365730				
ACCESSION	AX365730.1	GI:18697279			
KEYWORDS					
ORGANISM					
SOURCE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
ORIGIN					

Qy	1	ATGAGAGCCCGTCGCGCGCGCTCTCTGCGTCTCTGGTCTGTGAGCGCATCTCAAGGC	60
Db	67	ATGAGAGCCCTGCTGGCGCGCTCTCTCTCTGCGTCTCTGGTCTGTGAGCGCATCTCAAGGC	126
Qy	61	AGCATGAACTTCATCAAGTTCCATCGAATGATGCTGCTTAAATGAGAGAACTATGTG	120
Db	127	AGCATGAACTTCATCAAGTTCCATCGAATGATGCTGCTTAAATGAGAGAACTATGTG	186
Qy	121	TCCAACAAGTCTCTTCCAACTTCACTGTGTCAATCTGCCAAAGAAATTCGAGGGCAG	180
Db	187	TCCAACAAGTCTCTTCCAACTTCACTGTGTCAATCTGCCAAAGAAATTCGAGGGCAG	246
Qy	181	CACGTGAAATAGATAGTCAAAAACCTGTATAGAGGGAAATGTCACCTTTACCGAGA	240
Db	247	CACGTGAAATAGATAGTCAAAAACCTGTATAGAGGGAAATGTCACCTTTACCGAGA	306
Qy	241	AAGGCCAGACTGACACATGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	300
Db	307	AAGGCCAGACTGACACATGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	366
Qy	301	CAGAAACGATCAATGCGCAACAGATGATGCTTCACTGTGGGCGCTGGGAAACATAT	360
Db	367	CAGAAACGATCAATGCGCAACAGATGATGCTTCACTGTGGGCGCTGGGAAACATAT	426
Qy	361	TACTGACGAAACCCAGCAACCGGAGGCGACCTGTGTATGTGTGAGGTGGGCTTAAAG	420
Db	427	TACTGACGAAACCCAGCAACCGGAGGCGACCTGTGTATGTGTGAGGTGGGCTTAAAG	486
Qy	421	CCGCTTGTCCAAAGATGATGTGATGATGATGATGATGATGATGATGATGATGATGATG	480
Db	487	CCGCTTGTCCAAAGATGATGTGATGATGATGATGATGATGATGATGATGATGATGATG	546
Qy	481	CCAGAAATTTAAATTTAGTGTGCGCAAAAGACTGTGAGGCCCGCTTTAAGATTAT	540
Db	547	CCAGAAATTTAAATTTAGTGTGCGCAAAAGACTGTGAGGCCCGCTTTAAGATTAT	606
Qy	541	GGGGAGAAATTCACACACATCCGAGAACCGAGCTGTGTGGCGCATCTACAGAGGAC	600
Db	607	GGGGAGAAATTCACACACATCCGAGAACCGAGCTGTGTGGCGCATCTACAGAGGAC	666
Qy	601	CGGGGGGGCTGTGTCACTACGTGTGTGTGAGGACAGCTCATGAGCCCTTGTGGGTATC	660
Db	667	CGGGGGGGCTGTGTCACTACGTGTGTGTGAGGACAGCTCATGAGCCCTTGTGGGTATC	726
Qy	661	AGGCCACACACTGCTTCTATTATTATCCCAAGAGAGGAGACTATCATGCTTACCTGGT	720
Db	727	AGGCCACACACTGCTTCTATTATTATCCCAAGAGAGGAGACTATCATGCTTACCTGGT	786
Qy	721	CGCTCAAGCTTAACTCCCAACGCAAGAGAGGAGTGAATTTGAGTGGAAACCTATATC	780
Db	787	CGCTCAAGCTTAACTCCCAACGCAAGAGAGGAGTGAATTTGAGTGGAAACCTATATC	846
Qy	781	CTACACAAGACTTACAGCGCTGACACGCTTGTCTACCAACAAGCATTTGCTTGTGAAG	840
Db	847	CTACACAAGACTTACAGCGCTGACACGCTTGTCTACCAACAAGCATTTGCTTGTGAAG	906
Qy	841	ATCCGTTCCAAAGAGAGGAGGTGTGGAGAGCCATCCGAGACTTATACAGACCATTTGCTG	900
Db	907	ATCCGTTCCAAAGAGAGGAGGTGTGGAGAGCCATCCGAGACTTATACAGACCATTTGCTG	966
Qy	901	CCCTGATGTATATACGATCCCAAGTTTGGACAAGCTGTGATCACTGGCTTTGAGAAA	960
Db	967	CCCTGATGTATATATAGATCCCAAGTTTGGACAAGCTGTGATCACTGGCTTTGAGAAA	1026
Qy	961	GAGAAATTTCAACGACTATATCTTATCCGAGACGCTGAATAATGACTGTTGTGAAGCTGAT	1020
Db	1027	GAGAAATTTCAACGACTATATCTTATCCGAGACGCTGAATAATGACTGTTGTGAAGCTGAT	1086
Qy	1021	TCCCAACGGGAGGTGTACAGAGCCCACTACTACGCTGTGAAGTCAACCAAAATGCTG	1080

Db 1087 TCCACCGGGAGTGTGACGAGCCCACTACTAGGCTGTGAAGTCAACCAACCAATATCTG 1146
Qy 1081 TGTGCTGCTGACCAACAGTGTGAAGAAAGATTCTCTCCAGAGGAGCTAGAGGGAGCCCTTC 1140
Db 1147 TGTGCTGCTGACCAACAGTGTGAAGAAAGATTCTCTCCAGAGGAGCTAGAGGGAGCCCTTC 1206
Qy 1141 GTCTGTTCCCTCCAGAGCCGATGATCTTGAAGTGTGAAGTGTGAGCTGGGGCCGTGAGTGT 1200
Db 1207 GTCTGTTCCCTCCAGAGCCGATGATCTTGAAGTGTGAGCTGGGGCCGTGAGTGT 1266
Qy 1201 GCCCTGAGGACAGACGAGGCGTGTACAGAGAGTCTACACTTCTTACCTGTGATCCG 1260
Db 1267 GCCCTGAGGACAGACGAGGCGTGTACAGAGAGTCTACACTTCTTACCTGTGATCCG 1326
Qy 1261 AGTCACACCAAGAGAGAAATGCGCTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGAAATGCGCTGGCCCTCTGA 1362

RESULT 10
LOCUS BD226055 2294 bp DNA linear PAT 17-JUL-2003
DEFINITION Compound and method for remedy and diagnosis of lung cancer.
ACCESSION BD226055
VERSION BD226055.1 GI:33035825
KEYWORDS JP 2002533056-A/120.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 2294)
Reed,S.G. and Wang,T.
Compound and method for remedy and diagnosis of lung cancer
Patent: JP 2002533056-A 120 08-OCT-2002;
CORIXA CORP

COMMENT
OS Homo sapiens (human)
PN JP 2002533056-A/120
PD 08-OCT-2002
PF 17-MAR-1999 JP 20005336857
PR 18-MAR-1998 US 09/040802,18-MAR-1998 US 09/040984 PR
27-JUL-1998 US 09/123912,27-JUL-1998 US 09/123933 PI STEVEN
G REED, TONGTONG WANG
PC C12N15/09, A61K31/711, A61K35/14, A61K38/00, A61K38/16, A61K39/00,
A61K39/39,
A61K48/00, A61P35/00, C07K14/47, C07K16/18, C12N1/21, C12N5/10, PC
C1201/68,
PC G01N33/53, G01N33/574//C12P21/02, C12P21/08, C12N15/00, A61K37/02,
PC A61K37/04,
PC C12N5/00
CC Compound and method for remedy and diagnosis of lung cancer FH
Key Location/Qualifiers
FT source 1. 2294
FT Location/Qualifiers
1. 2294
Location/Qualifiers
1. 2294
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1. 2294
Location/Qualifiers
1. 2294
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1296; DB 6; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGAGCCTGTGCGCGCTCTCTGTGCGTCTGAGCGAGCTCCAAAGC 60
Db 67 ATGAGAGCCTGTGCGCGCTCTCTGTGCGTCTGAGCGAGCTCCAAAGC 126
Qy 61 AGCAATGACTTCAATCAAGTTCAATGAGTGTCTTAAATGAGAGAACTGTGT 120
Db 127 AGCAATGACTTCAATCAAGTTCAATGAGTGTCTTAAATGAGAGAACTGTGT 186

Qy 121 TCCAAAGAGTACTTCTCCAACTTCACTGTGTGCACTGCCCAAGAAATTCGAGGCGAG 180
Db 187 TCCAAAGAGTACTTCTCCAACTTCACTGTGTGCACTGCCCAAGAAATTCGAGGCGAG 246
Qy 181 CACTGTGAAATAGATAAGTCAAAAACCTGCTATGAGGGGAAATGTCACTTTACGAGA 240
Db 247 CACTGTGAAATAGATAAGTCAAAAACCTGCTATGAGGGGAAATGTCACTTTACGAGA 306
Qy 241 AAGGCGAGCTGACACACATGAGGCGCGGCTGCTGCTGCTGAACTTGTGCACTGTCTT 300
Db 307 AAGGCGAGCTGACACACATGAGGCGCGGCTGCTGCTGCTGAACTTGTGCACTGTCTT 366
Qy 301 CAGCAACGTAACCATGCGCCACAGATCTGATGCTCTTCACTGTGAGGCGCTGGGAAACATAT 360
Db 367 CAGCAACGTAACCATGCGCCACAGATCTGATGCTCTTCACTGTGAGGCGCTGGGAAACATAT 426
Qy 361 TACTGAGGAAACCCAGACAAACCGAGAGCGACCTGTGCTATGTGCAAGTGGGCTTAAG 420
Db 427 TACTGAGGAAACCCAGACAAACCGAGAGCGACCTGTGCTATGTGCAAGTGGGCTTAAG 486
Qy 421 CCGCTTGTCCAAAGTGCATGTGTGATGATGATGCGGAGATGGAAGAAAGCCCTCTCTCT 480
Db 487 CCGCTTGTCCAAAGTGCATGTGTGATGATGATGCGGAGATGGAAGAAAGCCCTCTCTCT 546
Qy 481 CCAAGAGATTTAAATTTAGTGTGCGCAAAAGCTGTGAGGCGCGCTTTAAGATTATT 540
Db 547 CCAAGAGATTTAAATTTAGTGTGCGCAAAAGCTGTGAGGCGCGCTTTAAGATTATT 606
Qy 541 GGGGAGAAATTTCAACCAATGAGAAACCAACCCCTGTGTTGCGGCACTTACAGAGGCGAC 600
Db 607 GGGGAGAAATTTCAACCAATGAGAAACCAACCCCTGTGTTGCGGCACTTACAGAGGCGAC 666
Qy 601 GGGGAGGCGCTGTGTCACTTACGTGTGTGTGAGGCGACCTTATCAAGCCTTGTGTGTATC 660
Db 667 GGGGAGGCGCTGTGTCACTTACGTGTGTGTGAGGCGACCTTATCAAGCCTTGTGTGTATC 726
Qy 661 AGGCGCACACAGCTTCACTTATTTACCAAGAGAGGAGACTACATGCTTACCTGGGT 720
Db 727 AGGCGCACACAGCTTCACTTATTTACCAAGAGAGGAGACTACATGCTTACCTGGGT 786
Qy 721 CGCTCAAGGCTTAACTTCAACAGCAGCAAGGAGAGTGAAGTTGAGTGAAGAACTTATC 780
Db 787 CGCTCAAGGCTTAACTTCAACAGCAGCAAGGAGAGTGAAGTTGAGTGAAGAACTTATC 846
Qy 781 CTACACAGAGCTACAGCGCTGACAGCGCTTGTCTACCAACAGCATTTGCTTGTGAAG 840
Db 847 CTACACAGAGCTACAGCGCTGACAGCGCTTGTCTACCAACAGCATTTGCTTGTGAAG 906
Qy 841 ATCCGTTCCAAAGAGGAGAGGAGTGTGCGGAGCCATCCCGGACTATACAGACATCTGCTG 900
Db 907 ATCCGTTCCAAAGAGGAGAGGAGTGTGCGGAGCCATCCCGGACTATACAGACATCTGCTG 966
Qy 901 CCGTCAATGATATACGATCCCAAGTTTGGAGCAAGCTGTGAGATCACTGTGCTTGGAAAA 960
Db 967 CCGTCAATGATATACGATCCCAAGTTTGGAGCAAGCTGTGAGATCACTGTGCTTGGAAAA 1026
Qy 961 GAGAAATTTACCACTATCTTATTCGAGAGCTGAAATTAAGTGTGTAAGCTGATT 1020
Db 1027 GAGAAATTTACCACTATCTTATTCGAGAGCTGAAATTAAGTGTGTAAGCTGATT 1086
Qy 1021 TCCACCGGAGAGTGTGAGGAGCCCACTACTAGGCTGTGAAGTCAACCAAAATGCTG 1080
Db 1087 TCCACCGGAGAGTGTGAGGAGCCCACTACTAGGCTGTGAAGTCAACCAAAATGCTG 1146
Qy 1081 TGTGCTGCTGACCAACAGTGTGAAGAAAGATTCTCTGCGAGGAGACTCAAGGGAGCCCTTC 1140
Db 1147 TGTGCTGCTGACCAACAGTGTGAAGAAAGATTCTCTGCGAGGAGACTCAAGGGAGCCCTTC 1206
Qy 1141 GTCTGTTCCCTCCAGAGCCGATGATCTTGAAGTGTGAGTGTGAGCTGGGGCCGTGAGTGT 1200
Db 1207 GTCTGTTCCCTCCAGAGCCGATGATCTTGAAGTGTGAGTGTGAGCTGGGGCCGTGAGTGT 1266
Qy 1201 GCCCTGAGGACAGAGCGCTTACAGAGAGTCTCAACTTCTTAACTTGGATCCG 1260

Db 1267 GCCCTGAAGGACACAGCCGCTCTACAGAGTCTACACTTCTTACCTTGATCCG 1326
Qy 1261 AGTCACACCAAGAGAGATGCGCTGCGCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGATGCGCTGCGCTCTGA 1362

RESULT 11
HMMUKA 2294 bp mRNA linear PRI 17-FEB-1998
LOCUS Homo sapiens mRNA for pro-urokinase precursor, complete cds.
DEFINITION D00244
ACCESSION D00244.1 GI:220138
VERSION pro-urokinase precursor, pro-urokinase.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Verde, P., Stoppelli, M.P., Galeffi, P., Di Nocera, P. and Blasi, P.
TITLE Identification and primary sequence of an unspliced human urokinase
poly(A)+ RNA
Proc. Natl. Acad. Sci. U.S.A. 81 (15), 4727-4731 (1984)
JOURNAL 8472706
MEDLINE 6589620
PUBMED 2 (sites)
REFERENCE Nagai, M., Hiramatsu, R., Kaneda, T., Hayasuke, N., Arimura, H.,
AUTHORS Nishida, M. and Suyama, T.
TITLE Molecular cloning of cDNA coding for human preprourokinase
JOURNAL Gene 36 (1-2), 183-188 (1985)
MEDLINE 86056954
PUBMED 2415429
REFERENCE 3 (bases 1 to 2294)
AUTHORS Hibino, Y., Miyake, T., Kobayashi, Y., Ohmori, M., Miki, T.,
TITLE Enhanced expression of human pro-urokinase cDNA in Escherichia coli
JOURNAL Agric. Biol. Chem. 52, 329-336 (1988)
FEATURES
source
1..2294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pPB3, pKT06, pKYU21"
/issue_type="kidney"
/note="2 bp upstream of PstI site"
67..1362
/codon_start=1
/product="pro-urokinase precursor"
/protein_id="BAA00175.1"
/db_xref="GI:220139"
/translation="MRALLARLLCVLVVSDSKSNELHQVPSNCDLNGTGVSNKY
FSNIHMCNCPKFGQHCIEDKSKTCYEGHGYRGKASDTMGRPCLPMNSATYVCO
TYAHRSDALOLGIGKHNVCNDRNRRPVCYVGLKPLVOCWHDADKRRSP
PEELKRCGCKTLPKRTIGRFTIENQWPAAYRNRHGSSTVYVGGSLISFCW
VIAATCFIDYPRKEDYIVLGRSLNNTQGMKFEVENLIHKQVSAADLHANNDI
ALKIKNSRCRCAQPSRTIQITCLPSMYNDPOFGTSCITGKENSSTYLYLPEOLKM
IVKLISHRECKQPHYYGSEVITKMLCAADPOWKTDSCGDSGGPLVCSLQGRMLTIG
IVSWGSCALKDKRPYTVRSHPLPWRSHTEKENGIAL"

CDS
67..126
/product="pro-urokinase"
127..1359
/product="pro-urokinase"
708
/citation=[1]
/replace="g"
843
/citation=[2]
/replace="c"
1065
/citation=[2]
/replace="g"
1146
/citation=[1]

sig_peptide
mat_peptide
conflict
conflict
conflict
conflict

conflict /replace="a"
1161 /citation=[1]
/replace="g"
conflict 1164 /citation=[1]
/replace="a"
polysignal 2259..2264
polysignal 2272..2277
ORIGIN
Query Match 100.0%; Score 1296; DB 9; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGAGCCCTGTGCGCGCCTGCTTCTCTGCTCTGCTGAGGACATCCAAAGC 60
Db 67 ATGAGAGCCCTGTGCGCGCCTGCTTCTCTGCTCTGCTGAGGACATCCAAAGC 126
Qy 61 AGCAATGAATCTCATCAAGTTCCATCGAACTGTGACTGTCTAAATGAGAAACATGTGTG 120
Db 127 AGCAATGAATCTCATCAAGTTCCATCGAACTGTGACTGTCTAAATGAGAAACATGTGTG 186
Qy 121 TCCAAAGTACTTCTCCAAACATTCAGTGTGCACTGCCAAAGAAATTCGAGAGGAG 180
Db 187 TCCAAAGTACTTCTCCAAACATTCAGTGTGCACTGCCAAAGAAATTCGAGAGGAG 246
Qy 181 CACTGTGAATAGATAGTCAAAAACCTGTATAGAGGGAAATGTCACTTTTACCGAGA 240
Db 247 CACTGTGAATAGATAGTCAAAAACCTGTATAGAGGGAAATGTCACTTTTACCGAGA 306
Qy 241 AAGGCGACATCGAACCAATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 307 AAGGCGACATCGAACCAATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
Qy 301 CAGCAAGGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 367 CAGCAAGGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
Qy 361 TACTGACAGAACCCGAGAACCGAGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 427 TACTGACAGAACCCGAGAACCGAGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
Qy 421 CGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 487 CGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
Qy 481 CCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
Db 547 CCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 606
Qy 541 GGGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
Db 607 GGGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 666
Qy 601 GGGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
Db 667 GGGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 726
Qy 661 AGCGCACACACTGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
Db 727 AGCGCACACACTGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 786
Qy 721 CGCTCAAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
Db 787 CGCTCAAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 846
Qy 781 CTACACAAGACTACAGGCTGACAGCGCTTCTACACCAACAGCAATTCGCTTCTGTAAG 840
Db 847 CTACACAAGACTACAGGCTGACAGCGCTTCTACACCAACAGCAATTCGCTTCTGTAAG 906
Qy 841 ATCCGTTCCAAAGAGGAGGCTGTGCGAGCCATCCCGAGCTATACAGACCATTCGCTG 900

Db 907 ATCCGTTCCAAGGAGGAGGAGTGTGCGAGCCATCCCGACTATACAGACCATCTGCTG 966
 QY 901 CCTTCATGATATACCATATCCCGAGTTTGGACAAAGCTGTGAGATCACTGGCTTTGAAAA 960
 Db 967 CCTTCATGATATACCATATCCCGAGTTTGGACAAAGCTGTGAGATCACTGGCTTTGAAAA 1026
 QY 961 GAGAAATTCACGAGTATCTATCCGAGAGCTGAAATGACTGTGTGAACTGAT 1020
 Db 1027 GAGAAATTCACGAGTATCTATCCGAGAGCTGAAATGACTGTGTGAACTGAT 1086
 QY 1021 TCCACCGGAGGTGTGACAGCCCACTACTAGGCTCTGAGTCAACCAAAATGCTG 1080
 Db 1087 TCCACCGGAGGTGTGACAGCCCACTACTAGGCTCTGAGTCAACCAAAATGCTG 1146
 QY 1081 TGTGCTGTGACCCCAAGTGAAGAAACGATTTCTGCGAGGAGACTAGGGGGACCCCTC 1140
 Db 1147 TGTGCTGTGACCCCAAGTGAAGAAACGATTTCTGCGAGGAGACTAGGGGGACCCCTC 1206
 QY 1141 GTCTGTTCCTCCAGAGCCGATGACTTGTGAGTGAATTTGAGCTGGGCGGTGAGTGT 1200
 Db 1207 GTCTGTTCCTCCAGAGCCGATGACTTGTGAGTGAATTTGAGCTGGGCGGTGAGTGT 1266
 QY 1201 GCCCTGAAGACAAGCCAGGCGTCTACAGAGTCTCACACTTCTTACCCTGGATCCG 1260
 Db 1267 GCCCTGAAGACAAGCCAGGCGTCTACAGAGTCTCACACTTCTTACCCTGGATCCG 1326
 QY 1261 AGTCACACCAAGAGAGATGGCTTGGCCCTCTGA 1296
 Db 1327 AGTCACACCAAGAGAGATGGCTTGGCCCTCTGA 1362

RESULT 12

G27040 2294 bp DNA linear STS 30-MAR-2000
 LOCUS SHGC-31374 Human Homo sapiens STS genomic, sequence tagged site.
 DEFINITION G27040
 ACCESSION G27040.1 GI:1375290
 VERSION STS.
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Buiakyoia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 JOURNAL 1 (bases 1 to 2294)
 COMMENT Unpublished, Olivier, M., Cox, D.R. (2000)
 Unpublished (2000)

Contact: Michael Olivier, David R. Cox
 Stanford Human Genome Center
 Stanford University School of Medicine
 4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
 Tel: (650) 320-5800
 Fax: (650) 320-5801
 Email: olivier@shgc.stanford.edu
 Primer A: ATCAGCTGTGAAGAGAGCTGGG
 Primer B: ATGCCCTGCCCTTTTAAT
 STS size: 225
 PCR Profile:
 Initial incubation: 94 degrees C for 90 seconds
 Denaturation: 94 degrees C for 15 seconds
 Annealing: 62 degrees C for 23 seconds
 Polymerization: 72 degrees C for 30 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600
 Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/uL
 Total Vol: 10 uL
 Buffer:

MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from D00244
 -- Washington University/Merck EST sequence.

FEATURES
 Source 1..2294
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="10"
 /clone_1b="Human"

STS
 primer_bind 1433..1454
 primer_bind 1433..1454
 complement (1438..1457)

ORIGIN

Query Match 100.0%; Score 1296; DB 11; Length 2294;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGGAGCGGCTGCTTCTGCGTCTGAGCGACTCCAAAGGC 60
 Db 67 ATGAGAGCCCTGCTGGAGCGGCTGCTTCTGCGTCTGAGCGACTCCAAAGGC 126
 QY 61 AGCAATGAATTGATCAAGTTCATGAACTGTGACTGTCTAAATGAGAGAACTGTG 120
 Db 127 AGCAATGAATTGATCAAGTTCATGAACTGTGACTGTCTAAATGAGAGAACTGTG 186
 QY 121 TCCAAAGTACTTCTCCAAATTCATCTGATGCACTGCGCCCAAGAAATTCGAGGCGAG 180
 Db 187 TCCAAAGTACTTCTCCAAATTCATCTGATGCACTGCGCCCAAGAAATTCGAGGCGAG 246
 QY 181 CACTGTGAATATAGTAAAGTCAAAACCTGTATGAGGAGAAATGCTCACTTTTACGAGGA 240
 Db 247 CACTGTGAATATAGTAAAGTCAAAACCTGTATGAGGAGAAATGCTCACTTTTACGAGGA 306
 QY 241 AAGGCCAGACCTGACACCAATGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 Db 307 AAGGCCAGACCTGACACCAATGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
 QY 301 CAGCAAACTGATCAATGCGCCCAAGATCTGATGCTTCTTCACTGCGGCTTGGGAAACATAAT 360
 Db 367 CAGCAAACTGATCAATGCGCCCAAGATCTGATGCTTCTTCACTGCGGCTTGGGAAACATAAT 426
 QY 361 TACTGAGGAAACCCAGACCAAGCGGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 427 TACTGAGGAAACCCAGACCAAGCGGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
 QY 421 CCGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 Db 487 CCGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
 QY 481 CCAGAAATTTAAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 Db 547 CCAGAAATTTAAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 606
 QY 541 GGGGAGAAATTCACCAATTCGAGACCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 Db 607 GGGGAGAAATTCACCAATTCGAGACCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
 QY 601 CCGGGGGGCTGTGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 Db 667 CCGGGGGGCTGTGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
 QY 661 AGGCGACACATGCTTCAATTTGATTTACCAAGAAAGAGAGACTCAATGCTTCTGCTGGGT 720
 Db 727 AGGCGACACATGCTTCAATTTGATTTACCAAGAAAGAGAGACTCAATGCTTCTGCTGGGT 786
 QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGGAGATGAAGTTTGAAGGTAACCTAATC 780

Db 787 CGCTCAAGGCTTAACCTCAACACGACGAGGGAGATGAATTGAGCTGAAAACTTAATC 846
QY 781 CTACACAGAGCTTAACGCGCTGACACGCTTGTCTACCAACAGCATTGCTTGTAAG 840
Db 847 CTACACAGAGCTTAACGCGCTGACACGCTTGTCTACCAACAGCATTGCTTGTAAG 906
QY 841 ATCCGTTCCAGAGGAGGAGGTGTGCGACGACATCCCGACCTATACACATCTGCTG 900
Db 907 ATCCGTTCCAGAGGAGGAGGTGTGCGACGACATCCCGACCTATACACATCTGCTG 966
QY 901 CCCTGAGATGATATGATGATCCCAAGTTTGACCAACAGCTGTGATCATCTGCTTGAAAA 960
Db 967 CCCTGAGATGATATGATGATCCCAAGTTTGACCAACAGCTGTGATCATCTGCTTGAAAA 1026
QY 961 GAGATTTCTACAGCACTATCTCTATCCGAGAGAGCTGAAAAATGACTGTGTGTAAGCTGATT 1020
Db 1027 GAGATTTCTACAGCACTATCTCTATCCGAGAGAGCTGAAAAATGACTGTGTGTAAGCTGATT 1086
QY 1021 TCCCAACCGGAGGTGTGACGACGCGCCACTACTAGCTGTGAAAGTCAACCAAAATGCTG 1080
Db 1087 TCCCAACCGGAGGTGTGACGACGCGCCACTACTAGCTGTGAAAGTCAACCAAAATGCTG 1146
QY 1081 TGTGTGCTGACCCCAAGTGTGAAAAAGATTCTGCGAGGAGACTCAAGGGAGACCCCTC 1140
Db 1147 TGTGTGCTGACCCCAAGTGTGAAAAAGATTCTGCGAGGAGACTCAAGGGAGACCCCTC 1206
QY 1141 GTCTGTCTCTCCAAAGGCGGATGACTTGTGAGTGAATGTGAGCTGTGGGCGCGTGATGT 1200
Db 1207 GTCTGTCTCTCCAAAGGCGGATGACTTGTGAGTGAATGTGAGCTGTGGGCGCGTGATGT 1266
QY 1201 GCCCTGAAAGGACCAAGCCAGGCGTCTTACACGAGAGCTTCAACTTTACCTGGATCCG 1260
Db 1267 GCCCTGAAAGGACCAAGCCAGGCGTCTTACACGAGAGCTTCAACTTTACCTGGATCCG 1326
QY 1261 AGTCACACCAAGAAAGAAATGGGCTGGCCCTCTGGA 1296
Db 1327 AGTCACACCAAGAAAGAAATGGGCTGGCCCTCTGGA 1362

RESULT 13
A35395
LOCUS A35395 2296 bp RNA linear PAT 11-DEC-1996
DEFINITION H.sapiens u-PA cDNA sequence.
ACCESSION A35395
VERSION A35395.1 GI:1926844
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eumetazoa; Euteleostomi; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2296)
AUTHORS Meyhack,B., Helm,U. and Buerigl,R.
TITLE Process for the production of proteins
JOURNAL Patent: EP 0288435-A 2 26-OCT-1988;
CIBA-GEIGY AG

FEATURES
source
location/Qualifiers
1..2296
/organism="Homo sapiens"
/mol_type="unassigned RNA"
/db_xref="taxon:9606"
70..1365
/gene="u-PA"
70..1365
/gene="u-PA"
/codon_start=1
/protein_id="CAA02215.1"
/db_xref="GI:1926845"
/translation="MRALLARLLCYLVVSDSKSGNEHLPVSNDCDCLNGSTCVSNKY
FSNIHWNCNPKKRGCGHCEIDSKTCTYBGNHFRYKASTDTMGSPCLPMSATVLAQ
TYHAHSDALQULGKHNYCNPNDRNRKPCVYGLRPLVECHVHCADGKSPSP
PEELKFCQCGKTLRPRFKYIGBFTTINQPFPAIYRHRGSGTVYVCGSLISPCW
VTSATHCFTDYPKEDYIVYLGSRSLNSNTQEMKFEVENLILHDDYASDVLAHNDI

ORIGIN
Query Match 99.9%; Score 1294.4; DB 6; Length 2296;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
ALIKRSKEGRCAQBSRTIQTILPSMYNDPQFGTSCEITGKENSITDYLPEQLKM
TVKLISHRECOOPHYGSEVTTKMLCAADPQMTKDSGQSGSLVLSQGRMTLTG
IVSMRGKALMKDKPGVYTRVSHFLPWISHTKEENGALL"

1 ATGAGAGCCCTGTGCGCGCGCTTCTCTGCGTCTGTGTGAGGAGCACTCCAAAAGC 60
70 ATGAGAGCCCTGTGCGCGCGCTTCTCTGCGTCTGTGTGAGGAGCACTCCAAAAGC 129
QY 61 AGCAATGAACCTTCATCAAGTTCCATGAACTGTGACTGTCTAATGAGAGCAATGTGTG 120
Db 130 AGCAATGAACCTTCATCAAGTTCCATGAACTGTGACTGTCTAATGAGAGCAATGTGTG 189
QY 121 TCCAAACAATCTTCTCCCAACTTCACTGTGCAACTGTGCACTGTGCACTGTGCACTGTG 180
Db 190 TCCAAACAATCTTCTCCCAACTTCACTGTGCAACTGTGCACTGTGCACTGTGCACTGTG 249
QY 181 CACTGTGAATATGATTAAGTCAAAACCTGCTATAGAGGGAATGTCATTTTACCGAGA 240
Db 250 CACTGTGAATATGATTAAGTCAAAACCTGCTATAGAGGGAATGTCATTTTACCGAGA 309
QY 241 AAGGCGAGACTGACACCATGGGCGCGCTGTGCTGCTGCTGCAACTGTGCTGCTGCTG 300
Db 310 AAGGCGAGACTGACACCATGGGCGCGCTGTGCTGCTGCTGCAACTGTGCTGCTGCTG 369
QY 301 CAGCAAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 370 CAGCAAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429
QY 361 TACTGACAGAAACCGACACCGAGGCGGACCTGTGCTATGTGACAGTGGGCTTAAG 420
Db 430 TACTGACAGAAACCGACACCGAGGCGGACCTGTGCTATGTGACAGTGGGCTTAAG 489
QY 421 CGGCTTGTCCAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 490 CGGCTTGTCCAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
QY 481 CCAGAAAGATTAATAATTTCACTGTGCGCAAAAGACTGTGAGGCGCGCTTAAAGATTA 540
Db 550 CCAGAAAGATTAATAATTTCACTGTGCGCAAAAGACTGTGAGGCGCGCTTAAAGATTA 609
QY 541 GGGGAGAAATTCACCACTATGAGAACGAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 610 GGGGAGAAATTCACCACTATGAGAACGAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 669
QY 601 CGGCGGCGCTGTGTCACTACCTAGTGTGAGGAGGAGCTCAACGCGCTTGTGAGTGTGATC 660
Db 670 CGGCGGCGCTGTGTCACTACCTAGTGTGAGGAGGAGCTCAACGCGCTTGTGAGTGTGATC 729
QY 661 AGCGGCACACACTGCTTCAATGATTAACCAAGAGAGAGACTACATGCTCACTGGGT 720
Db 730 AGCGGCACACACTGCTTCAATGATTAACCAAGAGAGAGACTACATGCTCACTGGGT 789
QY 721 CGCTCAAGGCTTAACCTCAACAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 790 CGCTCAAGGCTTAACCTCAACAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 849
QY 781 CTACACAGAGCTTACAGCGCTGACACGCTTGTCTACCAACAGCAATTTGCTTGGAAG 840
Db 850 CTACACAGAGCTTACAGCGCTGACACGCTTGTCTACCAACAGCAATTTGCTTGGAAG 909
QY 841 ATCCGTTCCAAAGGAGGAGGTGTGCGACGACATCCCGAGCTATATACAGACATCTGCTG 900
Db 910 ATCCGTTCCAAAGGAGGAGGTGTGCGACGACATCCCGAGCTATATACAGACATCTGCTG 969
QY 901 CCCTGAGATGATATGATGATCCCAAGTTTGACCAACAGCTGTGAGATCATCTGCTTGAAAA 960
Db 970 CCCTGAGATGATATGATGATCCCAAGTTTGACCAACAGCTGTGAGATCATCTGCTTGAAAA 1029

QY	961	GAGAAATTCACCGACTATCTCTATCTCCGAGACAGCTGAAATGAAGCTGTGTGAAGTGATT	1022
Db	1030	GAGAAATTCACCGACTATCTCTATCTCCGAGACAGCTGAAATGAAGCTGTGTGAAGTGATT	1089
QY	1021	TCCACCGGAGTGTACAGACCCCTACTACTACGGCTCTGAGATCACCAACCAATGCTG	1080
Db	1090	TCCACCGGAGTGTACAGACCCCTACTACTACGGCTCTGAGATCACCAACCAATGCTG	1149
QY	1081	TGTGCTGCTGACCCCACTGAGAAAACAGATTCTCTCCACGGAGACTCAGGGGACCCCTC	1140
Db	1150	TGTGCTGCTGACCCCACTGAGAAAACAGATTCTCTCCACGGAGACTCAGGGGACCCCTC	1209
QY	1141	GTCGTTCCTCCCAAGGCCGATGACTTGA CTGGAATTGTGAGCTGGGGCGTGGAGT	1200
Db	1210	GTCGTTCCTCCCAAGGCCGATGACTTGA CTGGAATTGTGAGCTGGGGCGTGGAGT	1269
QY	1201	GCCCTGGAAGACAAGCCAGCGCTACACGAGAGTCTCACACTTCTTACCTCTGATCCGC	1260
Db	1270	GCCCTGGAAGACAAGCCAGCGCTACACGAGAGTCTCACACTTCTTACCTCTGATCCGC	1329
QY	1261	AGTCACACCAAGAAAGAAATGGCTGGCCCTCTGTA	1296
Db	1330	AGTCACACCAAGAAAGAAATGGCTGGCCCTCTGTA	1365
RESULT 14			
BC013575		2333 bp mRNA linear PRI 08-OCT-2003	
LOCUS			
DEFINITION		Homo sapiens plasminogen activator, urokinase, mRNA (CDNA clone	
ACCESSION		MG:9223 IMAGE:3890980), complete ccd.	
VERSION		BC013575	
KEYWORDS		BC013575.2 GI:34194099	
SOURCE		MGC.	
ORGANISM		Homo sapiens (human)	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS		1 (bases 1 to 2333)	
		Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F., Ditchenko, L., Marusina, K., Farmer, A. A., Riddin, G. M., Hong, L., Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Schaefer, T. E., Brownstein, M. J., Udell, T. B., Tothiyaki, S., Carinci, P., Prange, C., Raha, S. S., Loguelfano, N. A., Peters, G. J., Abramson, R. D., Mullahy, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Wooley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakeley, R. W., Touchman, J. W., Green, R. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butcherfield, A. S., Krzywiński, M. I., Skalska, U. J., Smalls, D. E., Scherch, J., Schein, J. B., Jones, S. J., and Marz, M. A.	
TITLE		Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
MEDLINE		22388257	
PUBMED		12477932	
REFERENCE		2 (bases 1 to 2333)	
AUTHORS		Strausberg, R.	
TITLE		Direct Submission	
JOURNAL		Submitted (04-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomes Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK		NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT		On Aug 25, 2003 this sequence version replaced gi:15488888.	
		Contact: MGC help desk	
		Email: cgabbs-remail.nih.gov	
		Tissue Procurement: DCTD/DP/Gazdar	

DNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-sngc.stanford.edu>
 Contact: (Dickson, Mark) mdc@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/HLN! at: <http://image.llnl.gov>
 Series: IRAK P1.6: E. Consolium/HLN! at: <http://image.llnl.gov>
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505862.

Location/Qualifiers
 1..2333
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9223 IMAGE:3890980"
 /tissue_type="Lung, carcinoma, large cell undifferentiated."
 /clone_1lb="NIH MGC_69"
 /lab_host="DH10B"
 /note="vector: pCMV-Sport6"
 1..2333
 /gene="PLAU"
 /note="synonyms: URK, UPA"
 /db_xref="LocusID:5328"
 /db_xref="MIM:191840"
 93..1388
 /codon_start=1
 /product="plasminogen activator, urokinase"
 /protein_id="AAH13575.1"
 /db_xref="GI:15488889"
 /db_xref="LocusID:5328"
 /translation="MRLLRLRLLCVLVSDSKSNEIHOVPSNDCIANGCTCSNRYFSNHRSCPKRFGQCEIDKSTCYEGSGHLYRGASDTWGRPCLPWNASATVLQQTYSNHRSDALQIGLGRKNYCRNDNRNRPVQVQGLPLVQSCMHDCADGKPSPEELKFCQCKQTLRPKRIKIGESTFTENQPFPAIYRRRGSVTVYCGSILSPCVISATCFDIYPKKEDYIVYGLSRSLNSNGEMKEVENULIHKVYSATLVAHNDIALKIKRSGCQAPRSTIOTICLPASVYNDPQKTSCEITGPKENSTDYLYPQLKMTVYKLSHRECQGPHTYGSVTTKMLCADPQKTSQGSQSGPLVCSLQGRMTLTGIVSNGRSCALDKRPGVYTRSHPLPWIRSHTKENGAL"
 300..545
 /note="Kringle; Region: Kringle domain. Kringle domains have been found in plasminogen, heptocyte growth factors, prothrombin, and apolipoprotein A. Structure is disulfide-rich, nearly all-beta"
 /db_xref="CD:pfam0051"
 624..1349
 /note="Tryp. SPC; Region: Trypsin-like serine protease"
 /db_xref="CD:smart00020"

Query Match	99.9%	Score 1294.4	DB 9	Length 2333
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1295	Conservative 0	Mismatches 1	Indels 0	Gaps 0

```

OY      1  ATGAGAGCCCTGCTGGCGGCGCTGCTTCTCTGCGTCTTGCTCGTAGCGACTCCAAAGGC 60
Db      93  ATGAGAGCCCTGCTGGCGGCGCTGCTTCTCTGCGTCTTGCTCGTAGCGACTCCAAAGGC 152

OY      61  AGCAATGAACCTCATCAAGTTTCATCGCACTGGACCTGCTAATGAGAGAACTATGTTG 120
Db     153  AGCAATGAACCTTCATCAAGTTCCATCGCACTGGACCTGCTAATGAGAGAACTATGTTG 212

OY     121  TCCAAACAAGTACTTCTCCAACTTCACTGGTGGCACTGCCAAAGAAATTGCGAGGGCAG 180
Db     213  TCCAAACAAGTACTTCTCCAACTTCACTGGTGGCACTGCCAAAGAAATTGCGAGGGCAG 272

OY     181  CACTGTGAATAGATTAAGTCAAAAACCTGCTATGAGGGGAAATGTCACCTTTTACCGAGA 240
  
```

Db	273	CACGTGAATATAGTAATGATCAAAAACCTGGCTATAGAGGGAATGGTCACTTTTACCGAGGA	332
QY	241	AAGGCCAGCACTGACACCATATGGCCGGCCCTGCTGCCCTGGAACTTCCCATCTGTCCCTT	300
Db	333	AAGGCAGCACTGACACCATATGGCCGGCCCTGCTGCCCTGGAACTTCCCATCTGTCCCTT	392
QY	301	CAGCAAACTATACCATATGCCCAACAGATCTGATGCTTTTACGCTGGGCTGGGGAAACATAT	368
Db	393	CAGCAAACTATACCATATGCCCAACAGATCTGATGCTTTTACGCTGGGCTGGGGAAACATAT	452
QY	361	TACTGCAGGAACCCAGACAAACCGAGGCCGACCTGTGTCTATGTGACAGTGGGCTTAAG	420
Db	453	TACTGCAGGAACCCAGACAAACCGAGGCCGACCTGTGTCTATGTGACAGTGGGCTTAAG	512
QY	421	CCGCTGTTCAGAGTGATGGTGCATGACATGGGACATATGAAAAAAGCCCTCTCTCTCT	480
Db	513	CCGCTGTTCAGAGTGATGGTGCATGACATGGGACATATGAAAAAAGCCCTCTCTCTCT	572
QY	481	CCAGAAATTAATAATTTTCAGTGTGGCCAAAGAAGCTCTGAGGCCGCCCTTTAAGATTA	540
Db	573	CCAGAAATTAATAATTTTCAGTGTGGCCAAAGAAGCTCTGAGGCCGCCCTTTAAGATTA	632
QY	541	GGGGGGAATTAATCAACCATCTGGAAGAACACACCTCTGTTCGGCCATCTACAGAGGAC	600
Db	633	GGGGGGAATTAATCAACCATCTGGAAGAACACACCTCTGTTCGGCCATCTACAGAGGAC	692
QY	601	CGGGGGGGGCTCTGTCACTTACGATGTGGAGGAGAGCTCATCAAGCCCTTCTGGGTGATC	660
Db	693	CGGGGGGGGCTCTGTCACTTACGATGTGGAGGAGAGCTCATCAAGCCCTTCTGGGTGATC	752
QY	661	AAGGCCACACATCTGCTTCACTTATGATTACCAAGAGAGAGCATACATCGTCTACCTGGT	720
Db	753	AAGGCCACACATCTGCTTCACTTATGATTACCAAGAGAGAGCATACATCGTCTACCTGGT	812
QY	721	CGCTCAAGGCTTAACTTCCAAACGCAAGGGGAGATGAAGTTTATAGATGAAAACTTAATC	780
Db	813	CGCTCAAGGCTTAACTTCCAAACGCAAGGGGAGATGAAGTTTATAGATGAAAACTTAATC	872
QY	781	CTACACAGGACTACAGCGCTGACACGCTTGTCTACACACACAGCATTTGCTTGTCTGAAG	840
Db	873	CTACACAGGACTACAGCGCTGACACGCTTGTCTACACACACAGCATTTGCTTGTCTGAAG	932
QY	841	ATCCGTTCCAAAGGAGGGAGGTGTGGGACGCAATCCCGGACTTAACAGACATCTGCCCTG	900
Db	933	ATCCGTTCCAAAGGAGGGAGGTGTGGGACGCAATCCCGGACTTAACAGACATCTGCCCTG	992
QY	901	CCCTGGATATTAACGATCCCGAGTTTGGCACAGCTGTGAGATCACTAGGCTTTGGAAAA	960
Db	993	CCCTGGATATTAACGATCCCGAGTTTGGCACAGCTGTGAGATCACTAGGCTTTGGAAAA	1052
QY	961	GAGAAATTCATCCGACTATCTCTATCCGGACAGCTGAAAAATGACTGTGTGAAGCTGATT	1020
Db	1053	GAGAAATTCATCCGACTATCTCTATCCGGACAGCTGAAAAATGACTGTGTGAAGCTGATT	1112
QY	1021	TCCCAACCGGAGTGTACAGAGCCGCCACTTACCGGCTTGAAGTACCAACCAAAATGCTG	1080
Db	1113	TCCCAACCGGAGTGTACAGAGCCGCCACTTACCGGCTTGAAGTACCAACCAAAATGCTG	1172
QY	1081	TGTGTGCTGACCCACAGTGAAGAAAACAGATTCTGTCAAGGAGACTCAGAGGGGACCCCTC	1140
Db	1173	TGTGTGCTGACCCACAGTGAAGAAAACAGATTCTGTGTCAAGGAGACTCAGAGGGGACCCCTC	1232
QY	1141	GTCGTGTCCCTCCAAAGGCCGAGCTTGTGACATGGAATTGTATGAGCTGGGGCCGTGGATGT	1200
Db	1233	GTCGTGTCCCTCCAAAGGCCGAGCTTGTGACATGGAATTGTATGAGCTGGGGCCGTGGATGT	1292
QY	1201	GCCCTGAAGGACAAAGCCGATCTTACAGAGAGTCTCAACTTCTTAACCTGTGATCCGC	1260
Db	1293	GCCCTGAAGGACAAAGCCGATCTTACAGAGAGTCTCAACTTCTTAACCTGTGATCCGC	1352
QY	1261	AGTCACACCAAGGAGAGAAATGGCTGGCCCTCTCTGA	1296
Db	1353	AGTCACACCAAGGAGAGAAATGGCTGGCCCTCTCTGA	1388

RESULT 15				
E02577				
LOCUS	E02577	1296 bp	RNA	linear
DEFINITION	DNA encoding human prourokinase.			
ACCESSION	E02577			
VERSION	E02577.1	GI:2170807		
KEYWORDS	JP 1990227075-A/5.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1296)			
AUTHORS	Saeki,K., Nishi,T., Yasumura,S., Sato,M. and Ito,S.			
TITLE	NEW POLYPEPTIDE			
JOURNAL	Patent: JP 1990227075-A 5 10-SEP-1990;			
COMMENT	KYOMA HAKKO KOGYO CO LTD OS Homo sapiens (human)			

PF	28-SEP-1989 JP 1989253097	
PR	29-SEP-1988 JP 88P 245705	
PI	SASKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI SATO MORIYUKI,	
PI	ITOU SERGA	
PC	C12N9/72, C07K13/00, C07K15/14, C12N1/21, C12N5/10, C12N9/64, PC C12N15/27,	
PC	C12N15/58, C12N15/70, C12N15/85, C12P21/02, C12P21/02, (C12N1/21	
PC	C12R1:19),	
PC	(C12P21/02, C12R1:19), (C12P21/02, C12R1:91) ;	
CC	strandedness: Double,	
CC	topology: linear,	
CC	hypothetical: No;	
CC	anti-sense: No;	
CC	*source: cell type=Detroit 562 cell;	
CC	*source: clone=DVK11;	
PH	Key	Location/Qualifiers
PH		
FT	CDS	1. 1296
FT	/product='human prourokinase precursor'	FT
FT	61. 1293	
FT	/product='human peourokinase'.	

```
FEATURES
source      Location/Qualifiers
1..1296     /organism="Homo sapiens"
            /mol_type="genomic RNA"
            /db_xref="taxon:9606"
ORIGIN
```

Query Match	99.8%	Score 1292.8	DB 6	Length 1296	
Best Local Similarity	99.8%	Pred. No. 0			
Matches 1294	Conservative	0	Mismatches 2	Indels 0	Gaps 0
QY	1	ATGAAGGCCCTGCTGGCGCGCCTGCTTCTCTGCGCTCTGATGCGCATCTCCAAAGC	60		
Db	1	ATGAAGGCCCTGCTGGCGCGCGCTGCTTCTCTGGGTCTTGATGCGCATCTCCAAAGC	60		
QY	61	AGCAATGAACCTTCATCAAGTCCATCGAATGAGCATGCTCTAAATGAGAGAAATGATGTG	120		
Db	61	AGCAATGAACCTTCATCAAGTCCATCGAATGAGCATGCTCTAAATGAGAGAAATGATGTG	120		
QY	121	TCGAACAAAGTACTTCTCCAAACATTCACTGGTGCACATGCTCCAAAGAAATTGGAGGCGAG	180		
Db	121	TCGAACAAAGTACTTCTCCAAACATTCACTGGTGCACATGCTCCAAAGAAATTGGAGGCGAG	180		
QY	181	CACGTGTAATAGATTAAGTCAAAAAACGTGATAGAGGGAATGATGCATTTTAACGAGAGA	240		
Db	181	CACGTGTAATAGATTAAGTCAAAAAACGTGATAGAGGGAATGATGCATTTTAACGAGAGA	240		
QY	241	AAGGCAAGCACTGACACACATATGGCGCGCCTGTCTGCTCTGGAACTTGTCCATGTGCTTT	300		
Db	241	AAGGCAAGCACTGACACACATATGGCGCGCCTGTCTGCTCTGGAACTTGTCCATGTGCTTT	300		

QY 301 CAGCAACGTACCATGCCCCAGATCTGATGCTTTCAGCTGGGCTTGGGAAACATTAAT 360
DB 301 CAGCAACGTACCATGCCCCAGATCTGATGCTTTCAGCTGGGCTTGGGAAACATTAAT 360
QY 361 TACTGAGGAAACCCAGACAAACCGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 420
DB 361 TACTGAGGAAACCCAGACAAACCGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 420
QY 421 CCGCTGTGTCAGAGTGCATGTGTGATGACTGCGCAGATGGAAGAAAGCCCTCTCTCT 480
DB 421 CCGCTGTGTCAGAGTGCATGTGTGATGACTGCGCAGATGGAAGAAAGCCCTCTCTCT 480
QY 481 CCAAGAAATTAATTTTCAGTGTGGCCAAAGACTGAGGCCCGCTTAAGATTATT 540
DB 481 CCAAGAAATTAATTTTCAGTGTGGCCAAAGACTGAGGCCCGCTTAAGATTATT 540
QY 541 GGGGAGAAATTCACCAATCGAGAAACAGCCCTGTGTTGGGCACTACAGAGGAC 600
DB 541 GGGGAGAAATTCACCAATCGAGAAACAGCCCTGTGTTGGGCACTACAGAGGAC 600
QY 601 CCGGGGGGCTGTGTCACTACGTGTGTGAGGCAAGCTCATAGCCCTTGTGGTGATC 660
DB 601 CCGGGGGGCTGTGTCACTACGTGTGTGAGGCAAGCTCATAGCCCTTGTGGTGATC 660
QY 661 AGGCGACACACGTCTTCACTTGAATTAACCCAAAGAGAGACTACATCGTCTACCTGGGT 720
DB 661 AGGCGACACACGTCTTCACTTGAATTAACCCAAAGAGAGACTACATCGTCTACCTGGGT 720
QY 721 CCGTCAAGGCTTAATCTCAACAGCAAGGAGATGAATTTGAGGTGAAACTTAATC 780
DB 721 CCGTCAAGGCTTAATCTCAACAGCAAGGAGATGAATTTGAGGTGAAACTTAATC 780
QY 781 CTACACAAGACTACAGCGCTGACACGCTTGTCTACCAACAAGACATTTGCTTGAAG 840
DB 781 CTACACAAGACTACAGCGCTGACACGCTTGTCTACCAACAAGACATTTGCTTGAAG 840
QY 841 ATCCGTTCCAGAGAGGCGAGGTGTGGCGACCATCCCGAGCTATACAGACATCTGCTG 900
DB 841 ATCCGTTCCAGAGAGGCGAGGTGTGGCGACCATCCCGAGCTATACAGACATCTGCTG 900
QY 901 CCTCGATGTATACGATCCCAAGTTTGGCAACAAGCTGTGAGATCACTGCTTGGAAA 960
DB 901 CCTCGATGTATACGATCCCAAGTTTGGCAACAAGCTGTGAGATCACTGCTTGGAAA 960
QY 961 GAGAAATTTACCGACTATCTCTATCCGAGACAGCTGAAAATGACTGTTGAAAGCTGATT 1020
DB 961 GAGAAATTTACCGACTATCTCTATCCGAGACAGCTGAAAATGACTGTTGAAAGCTGATT 1020
QY 1021 TCCCAACCGGAGAGTGTGACGACGCCCACTATACGGCTCTGAAGTCAACCAAAATGCTG 1080
DB 1021 TCCCAACCGGAGAGTGTGACGACGCCCACTATACGGCTCTGAAGTCAACCAAAATGCTG 1080
QY 1081 TGTGCTGTGACCCAGAGTGTGAAACGATTTCTGCGAGGAGACTCAGGGGAGCCCTC 1140
DB 1081 TGTGCTGTGACCCAGAGTGTGAAACGATTTCTGCGAGGAGACTCAGGGGAGCCCTC 1140
QY 1141 GTCTGTTCCCTCCAAAGCCGCGATGACTTTGACTGGAATTGTGAGCTGGGGCCGATGAT 1200
DB 1141 GTCTGTTCCCTCCAAAGCCGCGATGACTTTGACTGGAATTGTGAGCTGGGGCCGATGAT 1200
QY 1201 GCGCTGAAGAGCAAGCCAGGCGTCTACAGAGTCTCACACTTTTAACTTGAATCCG 1260
DB 1201 GCGCTGAAGAGCAAGCCAGGCGTCTACAGAGTCTCACACTTTTAACTTGAATCCG 1260
QY 1261 AGTCACACCAAGAGAAATGAGCTGGCCCTCTGA 1296
DB 1261 AGTCACACCAAGAGAAATGAGCTGGCCCTCTGA 1296

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 15:42:28 ; Search time 600 Seconds
(without alignments)
9176.108 Million cell updates/sec

Title:	US-10-076-421-1
Perfect score:	1296
Sequence:	1 atgagagccctgtgcg.....agaatgacctgaccttga 1296

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
1: N_GeneSeqn_29Jan04:*
2: GeneSeqn1980s:*
3: GeneSeqn1990s:*
4: GeneSeqn2000s:*
5: GeneSeqn2001as:*
6: GeneSeqn2001bs:*
7: GeneSeqn2003as:*
8: GeneSeqn2003bs:*
9: GeneSeqn2003cs:*
10: GeneSeqn2004as:*
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1296	100.0	1296	6	ABA00207	Aba00207 gc-uPA ccc
2	1296	100.0	2281	2	AAZ24620	Aaz24620 Human ltr
3	1296	100.0	2294	3	AAc65859	Aac65859 Human ltr
4	1296	100.0	2294	6	ABl49078	AbI49078 Human ltr
5	1296	100.0	2294	6	ABQ92264	Abq92264 Human ltr
6	1296	100.0	2294	8	ADa28213	Ada28213 Human ltr
7	1296	100.0	2294	8	ADe53418	Ades53418 Human ltr
8	1294.4	99.9	2298	1	AAN80981	Aan80981 Sequence
9	1292.8	99.8	1296	2	AAQ06049	Aaq06049 plasmid F
10	1292.8	99.8	1296	2	AAQ05772	Aaq05772 Pro-uropki
11	1292.8	99.8	1473	1	AAN92037	Aan92037 Sequence
12	1292.8	99.8	1474	1	AAN50138	Aan50138 Sequence
13	1292.8	99.8	1475	2	AAZ24619	Aaz24619 Human ltr
14	1292.8	99.8	1475	3	AAc65858	Aac65858 Human ltr
15	1292.8	99.8	1475	6	ABZ35047	Abz35047 Human ger
16	1292.8	99.8	1475	6	ABl49077	AbI49077 Human ltr
17	1292.8	99.8	1475	6	ABQ92263	Abq92263 Human ltr
18	1292.8	99.8	1475	7	ACc78885	Acc78885 Human ltr
19	1292.8	99.8	1475	8	ADa28212	Ada28212 Human ltr
20	1292.8	99.8	1475	9	ADe53417	Ades53417 Human ltr
21	1292.8	99.8	1500	1	AAN91617	Aan91617 Sequence
22	1292.8	99.8	1964	4	AAH28220	Aah28220 Nucleotide
23	1292.8	99.8	1964	6	AAD27855	Aad27855 Human upR

ALIGNMENTS

24	1292.8	99.8	2299	1	AAN60703	Aan60703	Sequence
25	1292.8	99.8	2301	1	AAN70390	Aan70390	cDNA encod
26	1292.8	99.8	2301	1	AAN91075	Aan91075	DNA encod
27	1292.8	99.8	2336	6	ABZ35547	Abz35547	Human gen
28	1292.8	99.8	2336	6	ADB47346	Adb47346	Human cDN
29	1292.8	99.8	2346	6	ADe54011	Ades4011	Human gen
30	1292.8	99.8	2344	6	ABZ35272	Abz35272	Human gen
31	1291.2	99.6	1475	1	AAN50141	Aan50141	Sequence
32	1291.2	99.6	1475	1	AAN81558	Aan81558	Pro-UK at
33	1291.2	99.6	2328	9	ADB25630	Adb25630	Human cDN
34	1291.2	99.6	2341	8	ACH03958	Ach03958	Human cDN
35	1291.2	99.6	2377	2	AAQ020360	Aaq020360	Human pro
36	1291.2	99.6	2427	2	AAQ04107	Aaq04107	Human pro
37	1289.6	99.5	1296	6	ABE86598	Abes6598	Human p10
38	1289.6	99.5	1475	1	AAN70356	Aan70356	Human uri-
39	1289.6	99.5	1500	1	AAN91618	Aan91618	Human pro
40	1289.6	99.5	2303	2	AAQ73483	Aaq73483	Full lengt
41	1289.6	99.5	2304	6	AB159543	Ab159543	Human pro
42	1289.6	99.5	2304	6	ABX76437	Abx76437	Lung canc
43	1289.6	99.5	2304	7	ABX76275	Abx76275	Lung canc
44	1289.6	99.5	2304	7	ABX17681	Abx17681	DNA encod
45	1289.6	99.5	2304	7	ACP12920	Act12920	Human cer

RESULT 1

ABAU0207	ABA00207 standard; cDNA; 1296 BP.
ID	ABA00207
AC	ABA00207;
XX	
XX	15-NOV-2002 (first entry)
DT	
DE	sc-uPA coding sequence.
XX	
XX	Gene; single chain prepro-urokinase; sc-uPA; pro-urokinase; HIV;
KW	high molecular weight urokinase-type plasminogen activator; HMW-uPA;
KW	long A; long B; EGF-like domain; kringle domain; urokinase receptor;
KW	low molecular weight urokinase-type plasminogen activator; LMW-uPA; CD87
XX	binding domain; ss.
OS	Homo sapiens.
XX	
XX	Location/Qualifiers
PH	Key
FT	CDS
FT	1..1296
FT	/*tag= a
FT	/product= "sc-uPA"
FT	1..60
FT	/*tag= b
FT	61..1293
FT	mat_peptide
FT	/*tag= c
FT	/product= "pro-urokinase"
XX	
XX	EP1232755-A2.
PN	
PD	21-AUG-2002.
XX	
XX	15-FEB-2002; 2002EP-00003555.
PF	
XX	20-FEB-2001; 2001JP-00042655.
PR	19-JUN-2001; 2001JP-00184284.
XX	
PA	(JCRP-) JCR PHARM CO LTD.
XX	
XX	Wada M, Wada N;
PI	
XX	WPI; 2002-610512/66.
DR	P-PSDB; AAG79460.
XX	
XX	Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g. high
PT	molecular weight urokinase-type plasminogen activator, amino-terminal

PT fragment or an anti-CD87 antibody.

PS Disclosure: Page 17-20; 38pp; English.

XX This sequence encodes single chain prepro-urokinase (sc-uPA). Pro-urokinase (amino acids 21-431) with a cleavage between amino acid 178 and 179 gives high molecular weight urokinase-type plasminogen activator (HMW-uPA). HMW-uPA is a protein consisting of two peptide chains linked by a di-sulphide bond. The chains, long A and B, are formed by enzymatic cleavage between amino acids 178 and 179 of pro-urokinase. HMW-uPA includes an EGF-like domain, a kringle domain and a urokinase receptor (CD87) binding domain. HMW-uPA is then cleaved between amino acids 155 and 156 to give low molecular weight urokinase-type plasminogen activator (LMW-uPA) (amino acids 156-178 and 179-431), that has no plasminogen activator activity. sc-uPA, or fragments of it, may be used in the anti-HIV agents of the invention which comprise a ligand molecule that binds to CD87. The agents are useful for treating HIV-infected humans for suppression of reproduction of HIV. The anti-HIV agents act by a mechanism of action different from those of conventional drugs, widening the choice of therapeutic agents and avoiding problems of resistant HIV

XX Sequence 1296 BP; 328 A; 361 C; 337 G; 270 T; 0 U; 0 Other;

Query Match 100.0%; Score 1296; DB 6; Length 1296;

Best Local Similarity 100.0%; Pred. No. 2,9e-267;

Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAGAGCCCTGCTGGGGGCGGCTTCTCTGCGTCTGCTGCTGAGCGATCCAAAGC 60
DB 1 ATGAGAGCCCTGCTGGGGGCGGCTTCTCTGCGTCTGCTGCTGAGCGATCCAAAGC 60
QY 61 AGCAATGAATCTTCATCAAGTTCATGCACTGATGCTGATGCTGATGAGGAGGAGATG 120
DB 61 AGCAATGAATCTTCATCAAGTTCATGCACTGATGCTGATGCTGATGAGGAGGAGATG 120
QY 121 TCCAAACAAGTACTTCTCCAAATTCATGCTGAGCACTGCGCAAAAGAAATTCGAGGCG 180
DB 121 TCCAAACAAGTACTTCTCCAAATTCATGCTGAGCACTGCGCAAAAGAAATTCGAGGCG 180
QY 181 CACTGTGAATTAATAGTCAAAACCTGCTATGAGGGGAAATGTCATCTTTTACGAGGA 240
DB 181 CACTGTGAATTAATAGTCAAAACCTGCTATGAGGGGAAATGTCATCTTTTACGAGGA 240
QY 241 AAGGCGAGCACTGACACATGAGCGCGGCTGCTGCTGCTGAGCACTGCTGCACTGTC 300
DB 241 AAGGCGAGCACTGACACATGAGCGCGGCTGCTGCTGCTGAGCACTGCTGCACTGTC 300
QY 301 CAGCAAACTGACATGCGCCACAGATCTGATCTCTTCACTGAGGCTGAGGAAACATTA 360
DB 301 CAGCAAACTGACATGCGCCACAGATCTGATCTCTTCACTGAGGCTGAGGAAACATTA 360
QY 361 TACTGAGGAAACCGAGCAACCGGAGCGACCTGCTGCTATGAGGAGGAGGAGGAGGAG 420
DB 361 TACTGAGGAAACCGAGCAACCGGAGCGACCTGCTGCTATGAGGAGGAGGAGGAGGAG 420
QY 421 CCGCTGTCCAGAGTGCATGTCATGACTGCGCATGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 421 CCGCTGTCCAGAGTGCATGTCATGACTGCGCATGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 CCGAAGAAATTAATTTTCACTGCTGCGCAAAAGACTCTGAGGCGCGCTTTAAGATTAT 540
DB 481 CCGAAGAAATTAATTTTCACTGCTGCGCAAAAGACTCTGAGGCGCGCTTTAAGATTAT 540
QY 541 GGGGAGAAATTCACCAATCGAGAACGAGCCTGCTTGGCGGCACTCAACGAGGAGGAG 600
DB 541 GGGGAGAAATTCACCAATCGAGAACGAGCCTGCTTGGCGGCACTCAACGAGGAGGAG 600
QY 601 CGGGGGGGCTCTGCTCACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 601 CGGGGGGGCTCTGCTCACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 AAGGCGAGCACTGCTTCACTGATGTTACCAAAAGAGAGGAGGAGGAGGAGGAGGAGG 720
DB 661 AAGGCGAGCACTGCTTCACTGATGTTACCAAAAGAGAGGAGGAGGAGGAGGAGGAGG 720

```

```

DB 661 AAGGCGAGCACTGCTTCACTGATGTTACCAAAAGAGAGGAGGAGGAGGAGGAGGAGG 720
QY 721 CGCTCAAGGCTTAATCTCAACACGAGAGGAGATGAGTTGAGTGGAAAACTTAATC 780
DB 721 CGCTCAAGGCTTAATCTCAACACGAGAGGAGATGAGTTGAGTGGAAAACTTAATC 780
QY 781 CTACACAGAGATTAACGCGCTACACGCTTGTCTCCACACAGACATTTGCTGTGAAG 840
DB 781 CTACACAGAGATTAACGCGCTACACGCTTGTCTCCACACAGACATTTGCTGTGAAG 840
QY 841 ATCCGTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 841 ATCCGTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 CCTCGATGTATTAACGATCCCAAGTTTGGCAACAGCTGAGATCATCTGCTTGGAAAA 960
DB 901 CCTCGATGTATTAACGATCCCAAGTTTGGCAACAGCTGAGATCATCTGCTTGGAAAA 960
QY 961 GAGAAATTTTACCGAGATATCTATCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
DB 961 GAGAAATTTTACCGAGATATCTATCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
QY 1021 TCCCAACCGGAGTGTACAGAGCCCACTACTAGGCTCTGAGTCAACCAAAATGCTG 1080
DB 1021 TCCCAACCGGAGTGTACAGAGCCCACTACTAGGCTCTGAGTCAACCAAAATGCTG 1080
QY 1081 TGTGCTGCTGACCCACAGTGAAGAAACAGATTCTTCCAGGAGGAGGAGGAGGAGG 1140
DB 1081 TGTGCTGCTGACCCACAGTGAAGAAACAGATTCTTCCAGGAGGAGGAGGAGGAGG 1140
QY 1141 GTCTGTCCCTCCAAAGGCGGATGACTTGACTGGAATTGAGAGTGGGGGCGTGAATG 1200
DB 1141 GTCTGTCCCTCCAAAGGCGGATGACTTGACTGGAATTGAGAGTGGGGGCGTGAATG 1200
QY 1201 GCCCTGAAGGACAAAGCCAGGCGTCTTACAGAGAGTCTCACACTTCTTACCTGTAT 1260
DB 1201 GCCCTGAAGGACAAAGCCAGGCGTCTTACAGAGAGTCTCACACTTCTTACCTGTAT 1260
QY 1261 AGTCACACCAAGAAAGAAATGGCTGGCCCTCTGA 1296
DB 1261 AGTCACACCAAGAAAGAAATGGCTGGCCCTCTGA 1296

RESULT 2
AAZ24620
ID AAZ24620 standard; cDNA; 2281 BP.
XX
XX AAZ24620;
AC 07-DEC-1999 (first entry)
XX
XX DT
XX 07-DEC-1999 (first entry)
DE Human lung tumor associated polynucleotide.
XX
XX Human; lung tumor; lung cancer; T cell stimulation; ss.
KW Homo sapiens.
XX
XX PN
XX MO9947674-A2.
XX
XX PD
XX 23-SEP-1999.
XX
XX PF
XX 17-MAR-1999; 99WO-US005798.
XX
XX PR
XX 18-MAR-1998; 98US-00040802.
XX
XX PR
XX 18-MAR-1998; 98US-00040984.
XX
XX PR
XX 27-JUL-1998; 98US-00123912.
XX
XX PR
XX 27-JUL-1998; 98US-00123933.
XX
XX PA
XX (CORI-) CORIXA CORP.
PI Reed SG, Wang T;
XX
XX WPI, 1999-571839/48.

```

New isolated lung tumor polynucleotides, used to develop products for the treatment, prevention and monitoring the progression of lung cancer.

Claim 12; Page 115-116; 148pp; English.

The invention provides isolated human lung tumor nucleic acids and polypeptides. The polypeptides can be used for the treatment of lung cancer. The polypeptides and polynucleotides can be used to stimulate T cells or antigen presenting cells for use in the treatment of lung cancer. The polypeptides and monoclonal antibodies specific for the polypeptides can also be used to inhibit the development of lung cancer. Agents which bind the polypeptides can be used for detecting lung cancer and for monitoring the progression of lung cancer.

SQ Sequence 2281 BP; 549 A; 595 C; 604 G; 533 T; 0 U; 0 Other;

Query Match	100.0%	Score 1296;	DB 2;	Length 2281;
Best Local Similarity	100.0%;	Pred. No. 2.9e-267;		
Matches 1296;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGAGAGCCCTGTCGGCGGCGCTGCTTCTCTCGTCTCTGATCGTGAAGCACTCCAAGGC	60
Dp	67	ATGAGAGCCCTCTGGCGGCGCTGCTTCTCTCGTCTCTGATCGTGAAGCACTCCAAGGC	126
QY	61	AGCAATGAACCTTCATCAAGTTCATCGAACTGTGACTGTCTAATATGAGGAACATGTGTG	120
Dp	127	AGCAATGAACCTTCATCAAGTTCATCGAACTGTGACTGTCTAATATGAGGAACATGTGTG	186
QY	121	TTCCAAAGATTAAGTCTTCCCAACATTCACCTGGTGCACCTGCCCAAGAAATTCGGAGGGCAG	180
Dp	187	TTCCAAAGATTAAGTCTTCCCAACATTCACCTGGTGCACCTGCCCAAGAAATTCGGAGGGCAG	246
QY	181	CAGTGTGAATATGATATAGCTCAAAAACTGCTATATGAGGGGAATAGTCACTTTTACCAAGA	240
Dp	247	CAGTGTGAATATGATATAGCTCAAAAACTGCTATATGAGGGGAATAGTCACTTTTACCAAGA	306
QY	241	AAGGCAGACACTGACACCAATGGAGCGGCGCCCTGCTCCCTGGAACTCTGCACATGTCTTT	300
Dp	307	AAGGCAGACACTGACACCAATGGAGCGGCGCCCTGCTCCCTGGAACTCTGCACATGTCTTT	366
QY	301	CAGCAAAAGTACCATGCCCCACAGATCTATGTCTTCACTGGGCTCTGGGAAAACATAAT	360
Dp	367	CAGCAAAAGTACCATGCCCCACAGATCTATGTCTTCACTGGGCTCTGGGAAAACATAAT	426
QY	361	TACTGCAGAAACCCAGACACCGGAGGCGCACCTGGTGTATATGTGACAGGTGGGCTTAAG	420
Dp	427	TACTGCAGAAACCCAGACACCGGAGGCGCACCTGGTGTATATGTGACAGGTGGGCTTAAG	486
QY	421	CCGCTTGTCCAAAGTGCATGGTGCATGACTCGCAGATGAGAAAAAGCCCTTCTCTCTCT	480
Dp	487	CCGCTTGTCCAAAGTGCATGGTGCATGACTCGCAGATGAGAAAAAGCCCTTCTCTCTCT	546
QY	481	CCAGAGAAATTAATAATTTCACTGTGTGGCCAAAAGACTCTGAAGCCCGCTTTAAGATTAT	540
Dp	547	CCAGAGAAATTAATAATTTCACTGTGTGGCCAAAAGACTCTGAAGCCCGCTTTAAGATTAT	606
QY	541	GGGGGAGAAATTCACCAACCATGAGAAACAGGCCCTGGTTTGGCGGCATCTAAGAGAGGAC	600
Dp	607	GGGGGAGAAATTCACCAACCATGAGAAACAGGCCCTGGTTTGGCGGCATCTAAGAGAGGAC	666
QY	601	CGGGGGGGCTCTGTCACTTACCTGATGTGTGGAGGACGCTCATGAGCCCTTGTCTGGTGTATC	660
Dp	667	CGGGGGGGCTCTGTCACTTACCTGATGTGTGGAGGACGCTCATGAGCCCTTGTCTGGTGTATC	726
QY	661	AGCGGCACACACTGCTTCACTTATTAACCAAGAGAGAGACTACATCGTCTACCTGGGT	720
Dp	727	AGCGGCACACACTGCTTCACTTATTAACCAAGAGAGAGACTACATCGTCTACCTGGGT	786
QY	721	CGCTCAAGGCTTAATCCCAACACGGAAGGGGAGATGAATTTGAGGTGGAAAACTTAATC	780
Dp	787	CGCTCAAGGCTTAATCCCAACACGGAAGGGGAGATGAATTTGAGGTGGAAAACTTAATC	846

QY	781	CTACACAAAGGCTTACAGGCTGACACGTTGCTCCACACAAACGATTCCTTCCTAG	840
Db	847	CTACACAAAGGCTTACAGGCTGACACGTTGCTCCACACAAACGATTCCTTCCTAG	906
QY	841	ATCCGTTCCAAAGAGGGGAGGTGTGCCAGCATCCCGACTATACAGACCATCTGCCTG	900
Db	907	ATCCGTTCCAAAGAGGGGAGGTGTGCCAGCATCCCGACTATACAGACCATCTGCCTG	966
QY	901	CCCTCGATGTATACGATCCCACTTTGGCCAAAGCTGTGAGATCACTGGCTTTGAAAA	960
Db	967	CCCTCGATGTATACGATCCCACTTTGGCCAAAGCTGTGAGATCACTGGCTTTGAAAA	1026
QY	961	GAGAAATTCACCGACTATCTTATCCGAGGAGCTGAAAAATGACTGTTGTGAACCTGATT	1020
Db	1027	GAGAAATTCACCGACTATCTTATCCGAGGAGCTGAAAAATGACTGTTGTGAACCTGATT	1086
QY	1021	TCCACCGGGAGGTGTACAGACGCCCACTACAGGCTCTGAAATGTCAACCAACAAATCTG	1080
Db	1087	TCCACCGGGAGGTGTACAGACGCCCACTACAGGCTCTGAAATGTCAACCAACAAATCTG	1146
QY	1081	TGTCTCTGTACCCACAGTGGAAAAAGATTCCTGCCAGGGAGACTCAGGGGGACCCCTC	1140
Db	1147	TGTCTCTGTACCCACAGTGGAAAAAGATTCCTGCCAGGGAGACTCAGGGGGACCCCTC	1206
QY	1141	GTCGTTCCTCCCAAGGCGCGATACCTTTGACTGTGAATTGTGAACCTGGGCGGTGATGT	1200
Db	1207	GTCGTTCCTCCCAAGGCGCGATACCTTTGACTGTGAATTGTGAACCTGGGCGGTGATGT	1266
QY	1201	GCCTGAAAGCAAGCGAGGCGTCTACACGAGATCTCAACTTTCTTACCTGTATCGCG	1260
Db	1267	GCCTGAAAGCAAGCGAGGCGTCTACACGAGATCTCAACTTTCTTACCTGTATCGCG	1326
QY	1261	AGTCACACCAAGAGAGAAATGGCTGGCCCTCTGA	1296
Db	1327	AGTCACACCAAGAGAGAAATGGCTGGCCCTCTGA	1362

RESULT 3

ID AAC65859 standard; cDNA; 2294 BP.

AC AAC65859;

DT 21-FEB-2001 (first entry)

DE Human lung cancer-associated cDNA for contig 12.

KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;

XX

XX
:
9
2
1
9
9
9

XXXXXX

XX

PR 17-DEC-1999; 99US-00466396.

PR 10-JAN-2000; 2000US-00480884.

XX XX

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 104

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

XX

PT protein is used for

XX polynucleotides encoding lung tumor polypeptides, useful for treating
PT lung cancer or stimulating an immune response.
PT

PS Example 3; Page 232-233; 374pp; English.

CC The present invention describes human lung tumour proteins. Human lung
CC tumour proteins and polynucleotides have cytostatic and immunostimulant
CC activities, and can be used in vaccine production. Compositions
CC comprising the lung tumour proteins, polynucleotides, antibodies, fused
CC proteins, T cell populations, or antigen presenting cells that express
CC the lung tumour proteins are useful for treating lung cancer or
CC stimulating an immune response. ABL49859 to ABL49300 and ABL4946 to
CC ABL495070 represent sequences used in the exemplification of the present
CC invention

Sequence 2294 BP; 549 A; 596 C; 604 G; 545 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1296;	DB 6;	Length 2294;
-------------	---------	-------------	-------	--------------

Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	ATGAGAGCCCTGTGGCGGCGCTTCTCTGTGGTCTGGTGGAGCGACTCCAAGGC	60
Dp	67	ATGAGAGCCCTGTGGCGGCGCTTCTCTGTGGTCTGGTGGAGCGACTCCAAGGC	126
OY	61	AGCAATGAATTCATCAAGTTCCATGAACTGTGACTGTCTAATAAGAGGACATGTGTG	120
Dp	127	AGCAATGAATTCATCAAGTTCCATGAACTGTGACTGTCTAATAAGAGGACATGTGTG	186
OY	121	TTCACAAGTACTTCTCCAACTTCACTGTGTGCACTGCCCAAGAAATTGGAGGGCAG	180
Dp	187	TTCACAAGTACTTCTCCAACTTCACTGTGTGCACTGCCCAAGAAATTGGAGGGCAG	246
OY	181	CACGTGAAATGATPAAGTCAAAAACCTGTCTAATGAGGGGAATGTGACTTTTACCGAGA	240
Dp	247	CACGTGAAATGATPAAGTCAAAAACCTGTCTAATGAGGGGAATGTGACTTTTACCGAGA	306
OY	241	AAGGCGAGCACTGACACCACTGGGGCGGCGCTGCTCCCTGGAACCTCTGCACTGTCTT	300
Dp	307	AAGGCGAGCACTGACACCACTGGGGCGGCGCTGCTGCTGGAACCTCTGCACTGTCTT	366
OY	301	CAGCAAAAGTACCATGCCCCACAGATCTGATGCTTTCAAGCTGGGCTGTGGGAAACATAAT	360
Dp	367	CAGCAAAAGTACCATGCCCCACAGATCTGATGCTTTCAAGCTGGGCTGTGGGAAACATAAT	426
OY	361	TACTGACGAAGACCCGAGACACCTGGAGGGGCACTCTGTGTCTATGTGACAGGTGGGCTTAAG	420
Dp	427	TACTGACGAAGACCCGAGACACCTGGAGGGGCACTCTGTGTCTATGTGACAGGTGGGCTTAAG	486
OY	421	CCGCTTGTTCAGAGAGCGATGGTGTGACTGCTGGCAGATGGAAAAAAGCCCTCTCTCCCT	480
Dp	487	CCGCTTGTTCAGAGAGCGATGGTGTGACTGCTGGCAGATGGAAAAAAGCCCTCTCTCCCT	546
OY	481	CCAGAAGAATTAATAATTTCAAGTGTGGCCAAAAGACTCTGAGGCCCGGCTTTAAGATTATT	540
Dp	547	CCAGAAGAATTAATAATTTCAAGTGTGGCCAAAAGACTCTGAGGCCCGGCTTTAAGATTATT	606
OY	541	GGGGGGAATTAACAACAACATCGAAGAACCAAGCCCTGTGTTGGCGCCACTTCAAGAGAGCAC	600
Dp	607	GGGGGGAATTAACAACAACATCGAAGAACCAAGCCCTGTGTTGGCGCCACTTCAAGAGAGCAC	666
OY	601	CGGGGGGGGCTCTGTCAACCTACGATGTGTGAGGACAGCTCATCAAGCCCTTGTGGGTGATC	660
Dp	667	CGGGGGGGGCTCTGTCAACCTACGATGTGTGAGGACAGCTCATCAAGCCCTTGTGGGTGATC	726
OY	661	AGCGCCACACACTGTCTTCAATTGATTAACCCAAAGAGAGAGACTCAATCGCTTACCTGTGAT	720
Dp	727	AGCGCCACACACTGTCTTCAATTGATTAACCCAAAGAGAGAGACTCAATCGCTTACCTGTGAT	786
OY	721	CGCTCAAGGCTTAATCTCCAAACGCAAGGGAGATGAATTTGAGGTGGAAAACTTAATC	780
Dp	787	CGCTCAAGGCTTAATCTCCAAACGCAAGGGAGATGAATTTGAGGTGGAAAACTTAATC	846

QY	781	CTACGCAAGACTACAGCGCTTCAACAGCTTGTCTCAACCAACAACAAGCACTTGGCTTGTCAAG	84
Db	847	CTACGCAAGACTTACAGCGCTTCAACAGCTTGTCTCAACCAACAACAAGCACTTGGCTTGTCAAG	906
QY	841	ATCCGTTCCAAAGAGGAGGAGTGTGTGGAGCCATCCCGGACTTATACAGACCATCTGCGCTG	900
Db	907	ATCCGTTCCAAAGAGGAGGAGTGTGTGGAGCCATCCCGGACTTATACAGACCATCTGCGCTG	966
QY	901	CCCTCGATGTATATACGATCCCGAGTTTGGCACAGCTGTGAGTCACTGGCTTTGAAAA	966
Db	967	CCCTCGATGTATATACGATCCCGAGTTTGGCACAGCTGTGAGTCACTGGCTTTGAAAA	1020
QY	961	GAGAAATTCTACCGCACTATCTTATCCGAGAGCTGAAATATGACCTGTGGAAGCTGATT	1022
Db	1027	GAGAAATTCTACCGCACTATCTTATCCGAGAGCTGAAATATGACCTGTGTGAAGCTGATT	1080
QY	1021	TCCCAACGGGAGTGTACAGAGCCCACTACTACGCTCTGAAAGTCAACAACAAATAGCTG	1088
Db	1087	TCCCAACGGGAGTGTACAGAGCCCACTACTACGCTCTGAAAGTCAACAACAAATAGCTG	1144
QY	1081	TGTGTCTCTGACCCACAGTGA AAAACAGATTCTCTGCAGAGAGACTCAGGGGAGACCCCTC	1144
Db	1147	TGTGTCTCTGACCCACAGTGA AAAACAGATTCTCTGCAGAGAGACTCAGGGGAGACCCCTC	1200
QY	1141	GTCTGTCTCTCCCAAGGCGCGCATGTGACTGGAATTGTGAGCTGTGGGCGGTGAGTGT	1202
Db	1207	GTCTGTCTCTCCCAAGGCGCGCATGTGACTGGAATTGTGAGCTGTGGGCGGTGAGTGT	1268
QY	1201	GCCCTGAAGAGACAGCCAGGCGCTCTACAGAGAGTCTCACACTTCTTACCCTGATCCGC	1268
Db	1267	GCCCTGAAGAGACAGCCAGGCGCTCTACAGAGAGTCTCACACTTCTTACCCTGATCCGC	1322
QY	1261	AGTCACACCAAGAAAGAGATGCGCTGCGCTCTGA	1296
Db	1327	AGTCACACCAAGAAAGAGATGCGCTGCGCTCTGA	1362
RESULT 5			
ABQ92264	ABQ92264 standard; cDNA; 2294 BP.		
ID			
XX	AC	ABQ92264;	
XX	XX		
DT	07-OCT-2002	(first entry)	
XX	XX		
DE	Human lung cancer associated cDNA sequence SEQ ID NO:123.		
XX	XX		
KW	Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine; gene;		
XX	88.		
OS	Homo sapiens.		
XX	XX		
PN	WO200247534-A2.		
XX	XX		
PD	20-JUN-2002.		
XX	XX		
PF	30-NOV-2001; 2001WO-US047576.		
XX	XX		
PR	12-DEC-2000; 2000US-00735705.		
PR	07-MAY-2001; 2001US-00850716.		
XX	28-JUN-2001; 2001US-00897778.		
XX	XX		
PA	(CORI-) CORIXA CORP.		
XX	XX		
PI	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;		
PI	McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;		
PI	Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;		
XX	WPI; 2002-583465/62.		
XX			
PT	Novel lung carcinoma polynucleotide sequences and polypeptides encoded by		
PT	the polynucleotides, useful in pharmaceutical compositions such as		

PR 28-JUN-2000; 2000US-00606421.
 PR 02-AUG-2000; 2000US-00630940.
 PR 21-AUG-2000; 2000US-00643597.
 PR 15-SEP-2000; 2000US-00662786.
 PR 09-OCT-2000; 2000US-00685696.
 PR 12-DEC-2000; 2000US-00735705.
 PR 07-MAY-2001; 2001US-00850716.
 PR 28-JUN-2001; 2001US-00897778.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI Monnell PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
 XX
 DR WPI; 2003-540798/51.
 XX
 PT New isolated polynucleotides and polypeptides useful for diagnosing,
 PT preventing and/or treating cancer, particularly lung cancer.
 XX
 PS Example 3; Page 132-133; 296pp; English.
 XX
 CC The invention describes isolated polynucleotides and polypeptides useful
 CC for diagnosing, preventing and/or treating cancer, particularly lung
 CC cancer. A new isolated polynucleotide comprises: any of the 22 fully
 CC defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
 CC specification; complements of the nucleotide sequences cited above; at
 CC least 10 contiguous residues of the nucleotide sequences cited above; a
 CC sequence that hybridize to any of the nucleotide sequences under highly
 CC stringent conditions; a sequence that is at least 75 or 90% identical to
 CC the above nucleotide sequences; or degenerate variants of the above
 CC nucleotide sequences. The composition and methods are useful in
 CC diagnosing, preventing and/or treating cancer, particularly lung cancer,
 CC in gene therapy and in vaccines. This sequence represents a contig cDNA
 CC isolated from human lung tumour cells.
 SQ Sequence 2294 BP; 549 A; 596 C; 604 G; 545 T; 0 U; 0 Other:
 Query Match 100.0%; Score 1296; DB 8; Length 2294;
 Best Local Similarity 100.0%; Pred. No. 2.9e-267;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAGAGCCCTGCTGGCGCGCTGCTTCTGCGCTCGTGGAGCACTCCAAAGC 60
 DB 67 ATGAGAGCCCTGCTGGCGCGCTGCTTCTGCGCTCGTGGAGCACTCCAAAGC 126
 QY 61 AGCAATGAATTCATCAAGTTCCATCGAATCTGACTGTCTAAATGAGAGCAATGTGTG 120
 DB 127 AGCAATGAATTCATCAAGTTCCATCGAATCTGACTGTCTAAATGAGAGCAATGTGTG 186
 QY 121 TCCAAACAAGTACTTTCACCAATTCACCTGCTGCACTGCCCAAGAAATTCGAGGCGAG 180
 DB 187 TCCAAACAAGTACTTTCACCAATTCACCTGCTGCACTGCCCAAGAAATTCGAGGCGAG 246
 QY 181 CACTGTGAATATAGATAGTCAAAAACCTGCTATGAGGGGAATGTCACCTTTACCGAGGA 240
 DB 247 CACTGTGAATATAGATAGTCAAAAACCTGCTATGAGGGGAATGTCACCTTTACCGAGGA 306
 QY 241 AAGGCGAGCACTGACACCAATGAGCGCGCTGCTGCTGCACTGCACTGCTCTCT 300
 DB 307 AAGGCGAGCACTGACACCAATGAGCGCGCTGCTGCTGCACTGCACTGCTCTCTCT 366
 QY 301 CAGCAAAAGTACATGCTCCCAAGATCTGATGCTCTTCAAGTGGGCTGGGGAAACATAT 360
 DB 367 CAGCAAAAGTACATGCTCCCAAGATCTGATGCTCTTCAAGTGGGCTGGGGAAACATAT 426
 QY 361 TACTGAGAGAACCCAGAGAACCCGAGGCGACCTGCTGATGCTGAGGAGGCTCTTAAAG 420
 DB 427 TACTGAGAGAACCCAGAGAACCCGAGGCGACCTGCTGATGCTGAGGAGGCTCTTAAAG 486
 QY 421 CCGCTTGTCCAGAGTGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 487 CCGCTTGTCCAGAGTGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546

QY 481 CCAGAGAAATTAATTAATTTGAGTGGCCANAAGACTGTAGGGCCCGCTTTAAGATTATT 540
 DB 547 CCAGAGAAATTAATTAATTTGAGTGGCCANAAGACTGTAGGGCCCGCTTTAAGATTATT 606
 QY 541 GGGGGAGAAATTCACCAACATCGAGAACCAAGCCCTGTTTGGGCAATCTACAGAGGCGAC 600
 DB 607 GGGGGAGAAATTCACCAACATCGAGAACCAAGCCCTGTTTGGGCAATCTACAGAGGCGAC 666
 QY 601 CGGGGGGGGCTGTGTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 660
 DB 667 CGGGGGGGGCTGTGTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 726
 QY 661 AGGCGCACACACGCTTCAATTTATCCCAAAAGAGAGACTACATCCGCTACCTGCTGCT 720
 DB 727 AGGCGCACACACGCTTCAATTTATCCCAAAAGAGAGACTACATCCGCTACCTGCTGCT 786
 QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGATTTGAGGTGAGAAACCTTATC 780
 DB 787 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGATTTGAGGTGAGAAACCTTATC 846
 QY 781 CTACACAGAGACTACAGCGCTGACACGCTTGTCTACCAACAGACATTTGCTTGTGAG 840
 DB 847 CTACACAGAGACTACAGCGCTGACACGCTTGTCTACCAACAGACATTTGCTTGTGAG 906
 QY 841 ATCCGTTCCAAAGAGAGGCGAGTGTGGGCGACCTCCCGGCTATACAGACATCTGCTG 900
 DB 907 ATCCGTTCCAAAGAGAGGCGAGTGTGGGCGACCTCCCGGCTATACAGACATCTGCTG 966
 QY 901 CCTCATGATATACATGATCCCAAGTTGGGCAAGCTGTGATCACTGGCTTTGGAGAA 960
 DB 967 CCTCATGATATACATGATCCCAAGTTGGGCAAGCTGTGATCACTGGCTTTGGAGAA 1026
 QY 961 GAGAAATTCACGAGTATCTTATCCGAGACGCTGAAATGACTGTGTGAAGCTGATT 1020
 DB 1027 GAGAAATTCACGAGTATCTTATCCGAGACGCTGAAATGACTGTGTGAAGCTGATT 1086
 QY 1021 TCCCAACGGGAGGTGAGAGACCCCACTACTAGGCTCTGAAATCCCAACCAAAATGCTG 1080
 DB 1087 TCCCAACGGGAGGTGAGAGACCCCACTACTAGGCTCTGAAATCCCAACCAAAATGCTG 1146
 QY 1081 TGTGCTGTGACCCACAGATGAGAAACAGATTCTGTGCAGGGAGACTCAGGGGGAGCCCTC 1140
 DB 1147 TGTGCTGTGACCCACAGATGAGAAACAGATTCTGTGCAGGGAGACTCAGGGGGAGCCCTC 1206
 QY 1141 GTCTGTTCCCTCCAAAGCGCGCATGACTTGTGACTGAAATGTGAGCTGGGCGCTGAGATG 1200
 DB 1207 GTCTGTTCCCTCCAAAGCGCGCATGACTTGTGACTGAAATGTGAGCTGGGCGCTGAGATG 1266
 QY 1201 GCCCTGAAGAGCAAGCGCGCTTACAGAGAGTCTCACACTTTTACCTTGAATCCG 1260
 DB 1267 GCCCTGAAGAGCAAGCGCGCTTACAGAGAGTCTCACACTTTTACCTTGAATCCG 1326
 QY 1261 AGTCACACCAAGAGAGAGATGGCTGCGCCCTTGA 1296
 DB 1327 AGTCACACCAAGAGAGAGATGGCTGCGCCCTTGA 1362
 RESULT 7
 ADE53418
 ID ADE53418 standard; cDNA; 2294 BP.
 XX
 AC ADE53418;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human lung tumour protein cDNA sequence #120.
 XX
 KM Human; lung tumour protein; lung cancer; cytosolic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2003119763-A1.

XX 26-JUN-2003.
 PD 17-DEC-1999; 99US-00466396.
 XX 18-MAR-1998; 98US-00040802.
 PR 27-JUL-1998; 98US-00123912.
 PR 22-DEC-1998; 98US-00221107.
 PR 17-MAR-1999; 99WO-US005798.
 PR 02-APR-1999; 99US-00283479.
 XX (WANG/) WANG T.
 PA Wang T;
 PI WPI: 2003-874653/81.
 DR WPI: 2003-874653/81.
 XX Isolated pps comprising an immunogenic portion of a lung tumor protein
 PT used for the treatment and diagnosis of lung cancer.
 XX Claim 29; SEQ ID NO 123; 153pp; English.
 PS The present invention relates to the use of human lung tumor proteins,
 CC and the polynucleotide sequences that encode them. The invention also
 CC provides methods for inhibiting the development of lung cancer by the
 CC administration of pharmaceutical compositions or vaccines. A polypeptide
 CC comprising at least an immunogenic portion of a lung tumor protein, or a
 CC variant of the protein can be used in a pharmaceutical composition in the
 CC treatment and diagnosis of lung cancer. The polynucleotide sequences
 CC encoding the lung tumor proteins can also be used in the treatment and
 CC diagnosis of lung cancer. Prior art methods have been unable to detect
 CC early forms of lung cancer which this invention is able to provide. The
 CC present sequence represents a human lung tumor protein cDNA.
 XX
 SQ Sequence 2294 BP; 549 A; 596 C; 604 G; 545 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1296; DB 9; Length 2294;
 Best Local Similarity 100.0%; Pred. No. 2.9e-267;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 547 CCGAAGAAATTAAATTTCAGTGTGCGCAAAAGACTGAGGCCCGCGCTTTAAAGATTATTT 606
 QY 541 GGGGGAGAAATTCACCAACATCGAAGACCAAGCCCTGGTTGGCGCCATCTACAGAGGAC 600
 Db 607 GGGGGAGAAATTCACCAACATCGAAGACCAAGCCCTGGTTGGCGCCATCTACAGAGGAC 666
 QY 601 GGGGGGGGCTCTGTACCTTACGTTGTGAGGACGCTTATCAGCCCTTGTGGTGTATC 660
 Db 667 GGGGGGGGCTCTGTACCTTACGTTGTGAGGACGCTTATCAGCCCTTGTGGTGTATC 726
 QY 661 AGCGCCACACATGCTTATGATTACCCAAAGAGAGGACTCATGCTTACCTGGGT 720
 Db 727 AGCGCCACACATGCTTATGATTACCCAAAGAGAGGACTCATGCTTACCTGGGT 786
 QY 721 CGCTCAAGGCTTAACTCCCAACACGCAAGGGAGATGAAAGTTGAGTGAAGAAACCTAATC 780
 Db 787 CGCTCAAGGCTTAACTCCCAACACGCAAGGGAGATGAAAGTTGAGTGAAGAAACCTAATC 846
 QY 781 CTACACAAAGACTACAGGCGTGCACGCTTGTCTACCAACAGACATTGCTTGTGAAG 840
 Db 847 CTACACAAAGACTACAGGCGTGCACGCTTGTCTACCAACAGACATTGCTTGTGAAG 906
 QY 841 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGCCATCCGGAATTATACAGACCATGCGCTG 900
 Db 907 ATCCGTTCCAAAGAGGAGGAGTGTGCGAGCCATCCGGAATTATACAGACCATGCGCTG 966
 QY 901 CCTTGATGTATTAACGATCCCAAGTTTGGCAACAGCTGTGAGATCACTGGCTTGGAAA 960
 Db 967 CCTTGATGTATTAACGATCCCAAGTTTGGCAACAGCTGTGAGATCACTGGCTTGGAAA 1026
 QY 961 GAGAAATTCACGACTATCTATCCGGAAGAGCTGAAAATGACTGTGTAAGCTGATT 1020
 Db 1027 GAGAAATTCACGACTATCTATCCGGAAGAGCTGAAAATGACTGTGTAAGCTGATT 1086
 QY 1021 TCCCAACCGGAGTGTACAGACCCCACTACTACGAGCTCTGAAGTACACCAAAATGCTG 1080
 Db 1087 TCCCAACCGGAGTGTACAGACCCCACTACTACGAGCTCTGAAGTACACCAAAATGCTG 1146
 QY 1081 TGTGTCTGTACCCACAGTGTGAAAACAGATTCTTCCAGGAGACTACAGGGGAGCCCTTC 1140
 Db 1147 TGTGTCTGTACCCACAGTGTGAAAACAGATTCTTCCAGGAGACTACAGGGGAGCCCTTC 1206
 QY 1141 GTCTGTTCCCTCCAAAGGCGCATGACTTTGATGAAATTTGAGCTGAGGCGGTGATGT 1200
 Db 1207 GTCTGTTCCCTCCAAAGGCGCATGACTTTGATGAAATTTGAGCTGAGGCGGTGATGT 1266
 QY 1201 GCCCTGAAGACCAAGCGGCTTACACGAGAGTCTTCACTTTACCTTGATCCGC 1260
 Db 1267 GCCCTGAAGACCAAGCGGCTTACACGAGAGTCTTCACTTTACCTTGATCCGC 1326
 QY 1261 AGTCACACCAAGGAAGAAATGAGCTTGCTGCTCTGCA 1236
 Db 1327 AGTCACACCAAGGAAGAAATGAGCTTGCTGCTCTGCA 1362

RESULT 8
 AAN80981
 ID AAN80981 standard; DNA; 2298 BP.
 XX AAN80981;
 AC 25-MAR-2003 (revised)
 DT 14-SEP-1990 (first entry)
 XX Sequence of the single chain urokinase plasminogen activator (SCU-PA)
 DE cDNA insert prepared from human Hep3 cells.
 XX Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells;
 KW thrombosis treatment; thrombosis prevention; ss.
 XX OS Homo sapiens.
 XX


```

FH Key Location/Qualifiers
FT CDS 70..1365
FT /tag= a
FT /product= "peptide"
FT /tag= b
FT /tag= C
FT misc_feature 2298
FT /note= "polya site"
PN EP288435-A.
XX
XX 26-OCT-1988.
XX
XX 11-APR-1988; 88EP-00810234.
XX
XX 15-APR-1987; 87GB-00009081.
XX 16-JUN-1987; 87GB-00014059.
XX 04-DEC-1987; 87IE-00003299.
XX
XX (CIBA ) CIBA GEIGY AG.
XX
XX Meyhack B, Heim J, Burgi R;
XX
XX WPI: 1988-301440/43.
XX P-PSDB; AAP80430.
XX
XX Prod. of human single chain urokinase-type plasminogen activator - by
XX culturing yeast strain transformed with hybrid vector contg. yeast
XX expression control sequences.
XX
XX Example 1; Fig 2; 48bp; English.
XX
XX The patent is for the prodn. of human single chain urokinase-type
XX plasminogen activator (UTPA). Mutants of scu-PA are especially those
XX which render the protein protease resistant. The UTPA proteins exhibit
XX the biological activity of natural human UTPA without any refolding
XX procedure being necessary. They can be used as for known PAs in humans
XX for the prevention or treatment of thrombosis or other conditions where
XX it is desired to produce local fibrinolytic or proteolytic activity.
XX (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 2298 BP; 548 A; 594 C; 607 G; 548 T; 0 U; 1 Other:
XX
Query Match 99.9%; Score 1294.4; DB 1; Length 2298;
Best Local Similarity 99.9%; Pred. No. 6.4e-267;
Matches 1295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGAGCCCTGCTGGGGGCGCTGCTCTGCGTCTGCTGAGGCGAATCCAAAGGC 60
DB 70 ATGAGAGCCCTGCTGGGGGCGCTGCTCTGCGTCTGCTGAGGCGAATCCAAAGGC 129
QY 61 AGCAATGAATCTTCAATCAAGTTCCATCGAATGTAAGTCTTAAATGAGAGCAATGTGTG 120
DB 130 AGCAATGAATCTTCAATCAAGTTCCATCGAATGTAAGTCTTAAATGAGAGCAATGTGTG 189
QY 121 TCCAAACAAGTACTTTCACATCTTCACTGCTGCACTGCCCAAGAAATTCGAGGGCAG 180
DB 190 TCCAAACAAGTACTTTCACATCTTCACTGCTGCACTGCCCAAGAAATTCGAGGGCAG 249
QY 181 CACTGTGAATATAGTAAGTAAACCTGCTATGAGGGGAAATGCTACTTTTACCGAGGA 240
DB 250 CACTGTGAATATAGTAAGTAAACCTGCTATGAGGGGAAATGCTACTTTTACCGAGGA 309
QY 241 AAGGCGAGCATAGTACATGAGCGCGGCTGCTGCTGCAATCTTGCACTGTCTCT 300
DB 310 AAGGCGAGCATAGTACATGAGCGCGGCTGCTGCTGCAATCTTGCACTGTCTCT 369
QY 301 CAGCAAAACGTACATGCTCCACAGATCTGATGCTCTTCACTGGGCTGGGGAACATTAAT 360
DB 370 CAGCAAAACGTACATGCTCCACAGATCTGATGCTCTTCACTGGGCTGGGGAACATTAAT 429
QY 361 TACTGACAGAACCCAGACAAACCGGAGCGACCTGCTGCTATGTGACAGTGGGCTTAAAG 420
```

```

DB 430 TACTGACAGAACCCAGACAAACCGGAGCGACCTGCTGCTATGTGACAGTGGGCTTAAAG 489
QY 421 CCGCTTGTCCAAAGATGTCATGTGTCATGACTGCGCAGATGGAAGAAAGCCCTCTCTCT 480
DB 490 CCGCTTGTCCAAAGATGTCATGTGTCATGACTGCGCAGATGGAAGAAAGCCCTCTCTCT 549
QY 481 CCAAGAAATTAATTAATTTAGTGTGGCCAAAGACCTGTAGGGCCCGCTTAAAGATTAT 540
DB 550 CCAAGAAATTAATTAATTTAGTGTGGCCAAAGACCTGTAGGGCCCGCTTAAAGATTAT 609
QY 541 GGGGAGAAATTCACCAACATCGAAGAACCAAGCCCTGTTTGGGCACTTACAGAGGAC 600
DB 610 GGGGAGAAATTCACCAACATCGAAGAACCAAGCCCTGTTTGGGCACTTACAGAGGAC 669
QY 601 CGGGGGGGGCTGTGTCACTTACGTGTGTGAGAGGCAAGCTTCAACCCCTTGTGGTGATC 660
DB 670 CGGGGGGGGCTGTGTCACTTACGTGTGTGAGAGGCAAGCTTCAACCCCTTGTGGTGATC 729
QY 661 AGGCGCACACACTGCTTCAATGATTAACCAAGAGAGGACATACGCTTACCTGGGT 720
DB 730 AGGCGCACACACTGCTTCAATGATTAACCAAGAGAGGACATACGCTTACCTGGGT 789
QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAATTTGAGGTGGAAGAACTTATC 780
DB 790 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAATTTGAGGTGGAAGAACTTATC 849
QY 781 CTACACAGAGCTACACAGCGCTGACACGCTTGTCTACACACAGACATTCCTTCTGAAG 840
DB 850 CTACACAGAGCTACACAGCGCTGACACGCTTGTCTACACACAGACATTCCTTCTGAAG 909
QY 841 ATCCGTTTCCAAAGAGGCGAGGTGTGGCAGCCATCCCCGACTTATACAGACATTCCTG 900
DB 910 ATCCGTTTCCAAAGAGGCGAGGTGTGGCAGCCATCCCCGACTTATACAGACATTCCTG 969
QY 901 CCTCGATGATTAACGATCCCCAGTTTGGCAACAAGCTGTGATCACTGGCTTTGGAAAA 960
DB 970 CCTCGATGATTAACGATCCCCAGTTTGGCAACAAGCTGTGATCACTGGCTTTGGAAAA 1029
QY 961 GAGAAATTTACCGACTATCTCTATCCGAGACAGCTGAAATGACTGTTGTAAGCTGAT 1020
DB 1030 GAGAAATTTACCGACTATCTCTATCCGAGACAGCTGAAATGACTGTTGTAAGCTGAT 1089
QY 1021 TCCACCGGAGAGTGTGAGAGCCCGCACTACTAGGCTCTGAAATCAACCAAAATGCTG 1080
DB 1090 TCCACCGGAGAGTGTGAGAGCCCGCACTACTAGGCTCTGAAATCAACCAAAATGCTG 1149
QY 1081 TGTGCTGTGACCCACAGTGAAGAAACAGATTCTGTCAGGGAGACTCAAGGGAGACCCCT 1140
DB 1150 TGTGCTGTGACCCACAGTGAAGAAACAGATTCTGTCAGGGAGACTCAAGGGAGACCCCT 1209
QY 1141 GTCTGTTCCTCCAGAGCGGATGACTTGTGACTGAAATTTGTAGTGGGGCGGTGATGT 1200
DB 1210 GTCTGTTCCTCCAGAGCGGATGACTTGTGACTGAAATTTGTAGTGGGGCGGTGATGT 1269
QY 1201 GCGCTGAAGACAAAGCAGGCGTCTACAGAGTCTCACACTTTTAACTTGGATCCGC 1260
DB 1270 GCGCTGAAGACAAAGCAGGCGTCTACAGAGTCTCACACTTTTAACTTGGATCCGC 1329
QY 1261 AGTCACCAAGAGAGAGATGAGCTGGCCCTTGA 1296
DB 1330 AGTCACCAAGAGAGAGATGAGCTGGCCCTTGA 1365

```

RESULT 9
AA06049
ID AA06049 standard; DNA; 1296 BP.
XX
XX AA06049;
AC
XX
XX 25-MAR-2003 (revised)
DT 24-JAN-1991 (first entry)
XX

DE plasmid pUK1 pro-Urokinase sequence.
XX
KW pro-Urokinase; transgenic mice; ss.
XX
OS Synthetic.
XX
FN EP390592-A.
XX
PD 03-OCT-1990.
XX
PF 30-MAR-1990; 90EP-00303445.
XX
PR 31-MAR-1989; 89JP-00078574.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (EXPE-) CENT INST EXPER ANIMALS.
PA (JTK-) JIKKEN DOBUTSU CHUO KENK.
XX
PI Sekine S, Ito S, Katsuki M;
XX
DR WPI; 1990-299492/40.
DR P-PSDB; AAR07112.
XX
PT Prodn. of recombinant protein, esp. human pro:urokinase - from milk of
PT transgenic animals using promoter of bovine alpha S1 casein chromosomal
PT gene.
XX
PS Example; Table 1; 55pp; English.
XX
CC E.coli strain C600SF8 was transformed with recombinant plasmid containing
CC db DNA derived from human pharynx cancer cell strain Detroit 562. 10000
CC colonies were screened with the probe and one positive clone was
CC identified. Plasmid pUK1 was isolated and found to contain the coding
CC region and 3' non-coding region of pro-UK downstream of Cys(41). Four
CC silent substitutions were identified c.f. Holmes et al., Biotechnology,
CC vol.3, p.923 (1985) as follows: Asn(254), AAC to AAT; Leu(340), CTA to
CC CTG; Pro(345), CCC to CCA; Gln(346), CAA to CAG. See also AAO06045-Q06048
CC and AAO06392. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1296 BP; 327 A; 361 C; 337 G; 271 T; 0 U; 0 Other;
Query Match 99.8%; Score 1292.8; DB 2; Length 1296;
Best Local Similarity 99.8%; Pred. No. 1.4e-266;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGAGCCCTGTGCGCGCCCTGCTCTCTGCGCTCTGTCGAGCGAATCGCAAGGC 60
DB 1 ATGAGAGCCCTGTGCGCGCCCTGCTCTCTGCGCTCTGTCGAGCGAATCGCAAGGC 60
QY 61 AGCAATGAATTCATCAAGTTCCATCGAATGACTGTCTAATGAGAGACATGTGTG 120
DB 61 AGCAATGAATTCATCAAGTTCCATCGAATGACTGTCTAATGAGAGACATGTGTG 120
QY 121 TCCAAAGAGTACTTCTCCAACTTCCTGCTGCACTGCGCAAGAAATTCGAGGGCAG 180
DB 121 TCCAAAGAGTACTTCTCCAACTTCCTGCTGCACTGCGCAAGAAATTCGAGGGCAG 180
QY 181 CACTGTGAATATAGATAGTCAAAACCTGCTATGAGGGGAAATGTCATTTTACCGAGA 240
DB 181 CACTGTGAATATAGATAGTCAAAACCTGCTATGAGGGGAAATGTCATTTTACCGAGA 240
QY 241 AAGGCCGACCTGACACATGAGGCGGCGCTGCTGCGTGGAACTCTGCTGCTCTT 300
DB 241 AAGGCCGACCTGACACATGAGGCGGCGCTGCTGCGTGGAACTCTGCTGCTCTT 300
QY 301 CAGCAAACTTACCATGCGCCACATCTGATCTCTTCAAGTGGGCTTGAGGAAACATAAT 360
DB 301 CAGCAAACTTACCATGCGCCACATCTGATCTCTTCAAGTGGGCTTGAGGAAACATAAT 360
QY 361 TACTGACAGAACCCAGACAAACCGAGGCGAACCCTGTGTCTATGTGAGGTCGCTTAAG 420
DB 361 TACTGACAGAACCCAGACAAACCGAGGCGAACCCTGTGTCTATGTGAGGTCGCTTAAG 420

QY 421 CCGCTTGTCCAAAGATGCATGTGTCATGACTGCGCAATGAGAAAAAGCCCTCTCTCT 480
DB 421 CCGCTTGTCCAAAGATGCATGTGTCATGACTGCGCAATGAGAAAAAGCCCTCTCTCT 480
QY 481 CCAGAAAGATTTAAATTTTCAAGTGGGCCAAAAGACTCTGAGGCCCCGCTTTAAGTTATT 540
DB 481 CCAGAAAGATTTAAATTTTCAAGTGGGCCAAAAGACTCTGAGGCCCCGCTTTAAGTTATT 540
QY 541 GGGGGAGAAATTCACCAATCGAGAACAGCCCTGTTTGGCGCCATCTACAGAGGAC 600
DB 541 GGGGGAGAAATTCACCAATCGAGAACAGCCCTGTTTGGCGCCATCTACAGAGGAC 600
QY 601 CCGGGGGGCTCTGTCACTTACCTGTGTGAGAGGACCTTCATAGCCCTTGTGGGTATC 660
DB 601 CCGGGGGGCTCTGTCACTTACCTGTGTGAGAGGACCTTCATAGCCCTTGTGGGTATC 660
QY 661 AGCGCCACACATGCTCTTATTTATACCAAAAGAGAGAGACTATCATGTCTACCTGAGT 720
DB 661 AGCGCCACACATGCTCTTATTTATACCAAAAGAGAGAGACTATCATGTCTACCTGAGT 720
QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAAGTTTGAAGTGAAGAACTTAATC 780
DB 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAAGTTTGAAGTGAAGAACTTAATC 780
QY 781 CTACACAAAGACTACAGAGGCTGACACGCTTGCTACCAACAGCATGCTGCTGAG 840
DB 781 CTACACAAAGACTACAGAGGCTGACACGCTTGCTACCAACAGCATGCTGCTGAG 840
QY 841 ATCCGTTCCAAAGAGGAGGAGGTGTCGCGAGCCATCCCGAATTAACAGACCATGCTGCTG 900
DB 841 ATCCGTTCCAAAGAGGAGGAGGTGTCGCGAGCCATCCCGAATTAACAGACCATGCTGCTG 900
QY 901 CCTTGATGTATTAAGATATCCCAAGTTTGGCACAACTGTGATCATGCTTTGGAAAA 960
DB 901 CCTTGATGTATTAAGATATCCCAAGTTTGGCACAACTGTGATCATGCTTTGGAAAA 960
QY 961 GAGAAATTTACCGAATATCTCTATCCGAGGACGTGAAAAAGACTGTTGTGAAGCTGATT 1020
DB 961 GAGAAATTTACCGAATATCTCTATCCGAGGACGTGAAAAAGACTGTTGTGAAGCTGATT 1020
QY 1021 TCCACCGGAGAGTGTACAGACCCCACTAATGAGGCTCTGAAGTACCAACCAAAATGCTG 1080
DB 1021 TCCACCGGAGAGTGTACAGACCCCACTAATGAGGCTCTGAAGTACCAACCAAAATGCTG 1080
QY 1081 TGTGCTGTGACCCCAAGTGGAAAAAGATTCCTGCCAGAGGAGACTAGGGGGACCCCTC 1140
DB 1081 TGTGCTGTGACCCCAAGTGGAAAAAGATTCCTGCCAGAGGAGACTAGGGGGACCCCTC 1140
QY 1141 GTCTGTTCCTCCAAAGGCGGATGACTTTGAGTGAATTTGAGCTGGGCGCTGAGATGT 1200
DB 1141 GTCTGTTCCTCCAAAGGCGGATGACTTTGAGTGAATTTGAGCTGGGCGCTGAGATGT 1200
QY 1201 GCCCTGAAGGACAAAGCGGCTCTACAGAGATCTCAACTTTTACCTGGATCGC 1260
DB 1201 GCCCTGAAGGACAAAGCGGCTCTACAGAGATCTCAACTTTTACCTGGATCGC 1260
QY 1261 AGTCACCAAGAGAAATGAGCTGGCCCTCTGA 1296
DB 1261 AGTCACCAAGAGAAATGAGCTGGCCCTCTGA 1296
RESULT 10
AA055772
ID AA055772 standard; cDNA to mRNA; 1296 BP.
XX
AC AA055772;
XX
DT 13-JUN-1994 (first entry)
XX
DE Pro-urokinase derivative.
XX
KW Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic; factor;
KM 89.

OS	Homo sapiens.
XX	
XX	
FH	Location/Qualifiers
FT	1..1296
FT	/tag= a
FT	/product= "pro-urokinase_derivative"
XX	
PN	JP05336965-A.
XX	
PD	21-DEC-1993.
XX	
.PF	17-OCT-1991; 91JP-00269615.
XX	
PR	17-OCT-1991; 91JP-00269615.
XX	
PA	(KYOW) KYOMA HAKKO KOGYO KK.
XX	
DR	WPI; 1994-030907/04.
DR	P-PsDB; AAR47903.
PT	Novel human pro-urokinase derivs. having long half-life - with high thrombolytic activity, useful for treatment of thrombosis.
PS	Disclousure; Page 15-17; 29pp; Japanese.
CC	Sequences (AA055771-72) are pro-urokinase derivatives. The products have an inserted sugar moiety having an amino acid substituted, depleted or inserted variant around the thrombin cleavage site. They also have a long half-life allowing them to be used in the treatment of thrombosis
CC	
CC	
CC	
SQ	Sequence 1296 BP; 327 A; 361 C; 337 G; 271 T; 0 U; 0 Other;
Query Match	99.8%; Score 1292.8; DB 2; Length 1296;
Best Local Similarity	99.8%; Pred. No. 1.4e-266;
Matches 1294; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY	1 ATGAGAGCGCTGCCTGGCGGCGCTGTTCCTGCGCGTCTGATGAGGACTCCAAAGC 60
Db	1 ATGAAGCGCTGCTGCGGCGGCTGTTCCTGCGTCTGATGAGGACTCCAAAGC 60
OY	61 AGCATGAACTTCATCAAGTTCCATCGAATGTGACTGTCTTAATGAGGAACATGTGTG 120
Db	61 AGCATGAACTTCATCAAGTTCCATCGAATGTGACTGTCTTAATGAGGAACATGTGTG 120
OY	121 TCCAAAGAATCTTCCAACTTCACTGTGTCAAATGCGGCCAAGAAAATTGGAGGCGAG 180
Db	121 TCCAAAGAATCTTCCAACTTCACTGTGTCAAATGCGGCCAAGAAAATTGGAGGCGAG 180
OY	181 CACGTGAAATGATAATGATCAAAAAACGTATATGAGGGGAATGTCATTTTACCGAGA 240
Db	181 CACTGTAAATGATAATGATCAAAAAACGTATATGAGGGGAATGTCATTTTACCGAGA 240
OY	241 AAGGCACACATGACACCATTGGGCGGCGCTGCTGACCTGGAACTTGCCAATGTCTT 300
Db	241 AAGGCACACATGACACCATTGGGCGGCGCTGCTGACCTGGAACTTGCCAATGTCTT 300
OY	301 CAGCAAAAGTACATGATCCCAAGATCTGATGCTTTCAAGCTGGGCTTGAGGAAACAATAT 360
Db	301 CAGCAAAAGTACATGATCCCAAGATCTGATGCTTTCAAGCTGGGCTTGAGGAAACAATAT 360
OY	361 TACTGCAGAAACCCGAGAACCCGAGAGGGAGACCTGTGGTCTATATGTCAGAGTGGGCTTAAG 420
Db	361 TACTGCAGAAACCCGAGAACCCGAGAGGGAGACCTGTGGTCTATATGTCAGAGTGGGCTTAAG 420
OY	421 CGGCTTGTCCAAAGATGTGATGATGACTGCGACGAGATGAAAAAAGCCCTCTCTCTCT 480
Db	421 CGGCTTGTCCAAAGATGTGATGATGACTGCGACGAGATGAAAAAAGCCCTCTCTCTCT 480
OY	481 CCAGAAAGATTTAAATTTCAAGTGTGGCCAAAAGATCTTGAGGCCCCCGCTTTAAGATTAT 540
Db	481 CCAGAAAGATTTAAATTTCAAGTGTGGCCAAAAGATCTTGAGGCCCCCGCTTTAAGATTAT 540

QY	541	GGGGGAGAATTACACCATGTAGAACACAGCCCTGGTTGGGCATCTACAGAGGCAC	600		
Db	541	GGGGGAGATTTCACCAACCATGTAGAACACAGCCCTGGTTGGGCATCTACAGAGGCAC	600		
QY	601	CGGGGGGCTCTGTCACCTTACGTCGTGTGTAGAGGACCTCATCAGCCCTTGCTGGGTATC	660		
Db	601	CGGGGGGCTCTGTGTACCTTACGTCGTGTGTAGAGGACCTCATCAGCCCTTGCTGGGTATC	660		
QY	661	AGCGCCACACACTGCTTCATTGATTATCCAAAGAGAGAGACTATCATCTGCTACCTGGGT	720		
Db	661	AGCGCCACACACTGCTTCATTGATTATCCAAAGAGAGAGACTATCATCTGCTACCTGGGT	720		
QY	721	CGCTCAAGGCTTAACTCCAAACCCCAAGGGGAGATGAAAGTTTGGGTGAGAAACCTATATC	780		
Db	721	CGCTCAAGGCTTAACTCCAAACCCCAAGGGGAGATGAAAGTTTGGGTGAGAAACCTATATC	780		
QY	781	CTACACAAAGACTACAGGCTGACAGCCTTGCTCACCACAAACGACATTGGCTTGCTGAAG	840		
Db	781	CTACACAAAGACTACAGGCTGACAGCCTTGCTCACCACAAACGACATTGGCTTGCTGAAG	840		
QY	841	ATCCGTTTCCAAAGAGGGGAGGTGTGCGCAGCCATCCCGACCTATATACAGACCATCTGCTG	900		
Db	841	ATCCGTTTCCAAAGAGGGGAGGTGTGCGCAGCCATCCCGACCTATATACAGACCATCTGCTG	900		
QY	901	CCCTCGATGTATTACGATTCCTCCACTTTTGGCACAAGCTGTGAGATCATCTGGCTTTGGAAA	960		
Db	901	CCCTCGATGTATTACGATTCCTCCACTTTTGGCACAAGCTGTGAGATCATCTGGCTTTGGAAA	960		
QY	961	GAGAAATTCTACCGACTATCTCTATCCGGAGACGCTGAAAATGACTGTGTGTGAAGCTGATT	1020		
Db	961	GAGAAATTCTACCGACTATCTCTATCCGGAGACGCTGAAAATGACTGTGTGTGAAGCTGATT	1020		
QY	1021	TCCACCCGGGAGTGTACAGACCCCTCACTACCTACCGGCTCTGAAGTCAACACCAAAATGCTG	1080		
Db	1021	TCCACCCGGGAGTGTACAGACCCCTCACTACCTACCGGCTCTGAAGTCAACACCAAAATGCTG	1080		
QY	1081	TGTGTGTGTGACCCACAGTGTGAAAAAGATTCTCCGACGAGGAGACTCAGGGGGACCCCTTC	1140		
Db	1081	TGTGTGTGTGACCCACAGTGTGAAAAAGATTCTCCGACGAGGAGACTCAGGGGGACCCCTTC	1140		
QY	1141	GTCGTGTTCCCTCCAAAGGCGCGATGACTTTGACTGGAATTGTGAGCTGGGGCCGTTGAGTGT	1200		
Db	1141	GTCGTGTTCCCTCCAAAGGCGCGATGACTTTGACTGGAATTGTGAGCTGGGGCCGTTGAGTGT	1200		
QY	1201	GCCTGTGAAGACAAGCCAGGCGTCTACACGAGAGTCTCACACTTCTTACCTCGATTCGCG	1260		
Db	1201	GCCTGTGAAGACAAGCCAGGCGTCTACACGAGAGTCTCACACTTCTTACCTCGATTCGCG	1260		
QY	1261	AGTCACACCAAGAAAGAAATGCGCTTGCGTCTGA	1296		
Db	1261	AGTCACACCAAGAAAGAAATGCGCTTGCGTCTGA	1296		
RESULT 11					
ID	AA92037	standard; DNA; 1473 BP.			
XX	AA92037;				
AC	AA92037;				
XX	25-MAR-2003	(revised)			
DT	03-APR-1990	(first entry)			
XX	Sequence of variant human prourokinase.				
DE	Variant human prourokinase; plaamid pUH3; <i>Saccharomyces cerevisiae</i> AH22.				
XX	Homo sapiens.				
OS	Homo sapiens.				
XX	Key	Location/Qualifiers			
PH	CDS	81..1376			
FT	/*tag= a				
XX	JP01252283-A				

PT epithelial, and muscle cell types, useful for identifying disease
PT pathologies involving alterations of gene expression, e.g. cancer.
XX
XX
PS Claim 3, Page 382-383; 850pp; English.

XX
XX
CC The invention relates to a gene expression profile comprising one or more
CC genes (AB234889-AB235692) and generated from a cell type. The cell type
CC is a coronary artery endothelium, umbilical artery or vein endothelium,
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC endothelium, myometrium microvascular endothelium, keratinocyte
CC epithelium, bronchial epithelium, mammary epithelium, prostate
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC small airway epithelium, renal epithelium, umbilical artery smooth
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC osteoblasts or prostate stromal cell. The gene expression profile is
CC for determining the level of RNA expression for a sample, determining the
CC phenotype of a cell and distinguishing cell types. The gene or a protein
CC expression profile is useful in identifying disease pathologies involving
CC alterations of gene expression. The assessment of expression profiles may
CC provide meaningful information with respect to tumour type and stage,
CC treatment methods, and prognosis. The gene or protein expression profile
CC may also be used for creating microarrays. The microarray is useful for
CC genetic and physical mapping of genomes. DNA sequencing, genetic or
CC medical diagnosis, genotyping of organisms, confirming cell or tissue
CC identifications and in identifying promising antibiotics, antiviral or
CC antifungal agents.
XX
XX

Sequence 1475 BP; 357 A; 429 C; 386 G; 303 T; 0 U; 0 Other;

Query Match 99.8%; Score 1292.8; DB 6; Length 1475;
Best Local Similarity 99.8%; Pred. No. 1.4e-266;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGCGGCGCTGCTTCTGCGCTGCTGCGAGCTCCAAAGCG 60
DB 81 ATGAGAGCCCTGCTGCGGCGCTGCTTCTGCGCTGCTGCGAGCTCCAAAGCG 140
QY 61 AGCAATGAACTTCAATCAAGTTCAATGAACTGAACTGAACTGAACTGAACTG 120
DB 141 AGCAATGAACTTCAATCAAGTTCAATGAACTGAACTGAACTGAACTGAACTG 200
QY 121 TCCAAAGAACTTCTTCCAAATTCATCTGTCGAACTGCGCAAGAAATTCGAGGCG 180
DB 201 TCCAAAGAACTTCTTCCAAATTCATCTGTCGAACTGCGCAAGAAATTCGAGGCG 260
QY 181 CACTGTGAATTAAGTAAGTCAAAACCTGCTATGAGGGAATGTCACCTTTTACCGAGA 240
DB 261 CACTGTGAATTAAGTAAGTCAAAACCTGCTATGAGGGAATGTCACCTTTTACCGAGA 320
QY 241 AAGGCCAGCACTGACACCATGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 321 AAGGCCAGCACTGACACCATGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 380
QY 301 CAGCAAAAGTACATGAGCCCAAGATCTGATCTTCACTGAGGCTGAGGGAACATAAT 360
DB 381 CAGCAAAAGTACATGAGCCCAAGATCTGATCTTCACTGAGGCTGAGGGAACATAAT 440
QY 361 TACTGAGAGAACCCAGAACACCGGAGGAGACCTGCTGCTATGTCAGGTGAGGCTTAAAG 420
DB 441 TACTGAGAGAACCCAGAACACCGGAGGAGACCTGCTGCTATGTCAGGTGAGGCTTAAAG 500
QY 421 CCGCTTGTCCAAAGAGTGCATGTCATGTCGAGTCGAGTGAAGAAAGCCCTCTCTCT 480
DB 501 CCGCTTGTCCAAAGAGTGCATGTCATGTCGAGTCGAGTGAAGAAAGCCCTCTCTCT 560
QY 481 CCGAAGAAATTAATTTCACTGTCGCAAAAGACTTGAAGGCTTGAAGTATTT 540
DB 561 CCGAAGAAATTAATTTCACTGTCGCAAAAGACTTGAAGGCTTGAAGTATTT 620
QY 541 GGGGGAATTCACCAATGAGAACGAGCCCTGCTTGGGGCAATCTACAGAGGAC 600

DB 621 GGGGGAATTCACCAATGAGAACGAGCCCTGCTTGGGGCAATCTACAGAGGAC 680
QY 601 CCGGAGGAGTCTGTCACTGATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
DB 681 CCGGAGGAGTCTGTCACTGATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 740
QY 661 AGCGGAGGAGTCTGTCACTGATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 741 AGCGGAGGAGTCTGTCACTGATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 800
QY 721 CCGTGAAGCTTAACTCCAAACGCAAGGAGGAGTGAAGTTGAGTGAAGTGAATTC 780
DB 801 CCGTGAAGCTTAACTCCAAACGCAAGGAGGAGTGAAGTTGAGTGAAGTGAATTC 860
QY 781 CTACAGAGGAGTCAAGGAGTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
DB 861 CTACAGAGGAGTCAAGGAGTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 920
QY 841 ATCCGTTCCAAAGAGGAGGAGTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 900
DB 921 ATCCGTTCCAAAGAGGAGGAGTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 980
QY 901 CCGTGAAGCTTAACTCCAAACGCAAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 960
DB 981 CCGTGAAGCTTAACTCCAAACGCAAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1040
QY 961 GAGAAATTCACCAATGATCTATCCGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
DB 1041 GAGAAATTCACCAATGATCTATCCGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1100
QY 1021 TCCCAAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 1101 TCCCAAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1160
QY 1081 TGTGCTGTCACCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
DB 1161 TGTGCTGTCACCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1220
QY 1141 GTCTGTTCTCCCAAGGCGGATGATCTTGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1200
DB 1221 GTCTGTTCTCCCAAGGCGGATGATCTTGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1280
QY 1201 GCCCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
DB 1281 GCCCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1340
QY 1261 AGTCACACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
DB 1341 AGTCACACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1400

Search completed: March 21, 2004, 18:17:07
Job time : 610 secs

0 4 1 2 3 4 5 6 7 8 9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Db 307 AAGGCCAGCATGACACATGAGCCGCGCCCTGCTGCTCCCTGAACTCTGCCACTGTCCTT 366
QY 301 CAGCAAAAGCTACCTATGCCACAGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAT 360
Db 367 CAGCAAAAGCTACCTATGCCACAGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAT 426
QY 361 TACTGCAAGAACCCAGACACCGGAGGGGACCTTGTGCTATGTGCAAGTGGGCTTAAAG 420
Db 427 TACTGCAAGAACCCAGACACCGGAGGGGACCTTGTGCTATGTGCAAGTGGGCTTAAAG 486
QY 421 CGCTTGTTCAGAGTGTGATGTGATGCTGCGAGATGAGAAAAAGCCCTCTCTCT 480
Db 487 CGCTTGTTCAGAGTGTGATGTGATGCTGCGAGATGAGAAAAAGCCCTCTCTCT 546
QY 481 CCAGAAAGATTTAAATTTTCAGTGTGCGCAAAAGACTCTGAGGCCCTGTTAAGATTAT 540
Db 547 CCAGAAAGATTTAAATTTTCAGTGTGCGCAAAAGACTCTGAGGCCCTGTTAAGATTAT 606
QY 541 GGGGGAATTCACCAACCATGAGAACAGGCCCTGTGTTGGGCACTCTACAGAGGCAC 600
Db 607 GGGGGAATTCACCAACCATGAGAACAGGCCCTGTGTTGGGCACTCTACAGAGGCAC 666
QY 601 GGGGGGGCTCTGTCACTTACCTGTGTGTGAGAGCAGCCCTCATCAGCCCTTGTGGTATC 660
Db 667 GGGGGGGCTCTGTCACTTACCTGTGTGTGAGAGCAGCCCTCATCAGCCCTTGTGGTATC 726
QY 661 AGCGGCACACAGTGTTCATGATTTACCCAAAGAAAGAGACTCATCTGTCTACCTGGT 720
Db 727 AGCGGCACACAGTGTTCATGATTTACCCAAAGAAAGAGACTCATCTGTCTACCTGGT 786
QY 721 CGCTCAAGGCTTAACTCCAAACAGCAAGGGAGATGAAGTTTGAAGTGAAGAAACCTATC 780
Db 787 CGCTCAAGGCTTAACTCCAAACAGCAAGGGAGATGAAGTTTGAAGTGAAGAAACCTATC 846
QY 781 CTACAAAGAGCTACAGGCGTGAACAAGCTTGTCTACCAACAAGAGATTTGCTTGTGAAG 840
Db 847 CTACAAAGAGCTACAGGCGTGAACAAGCTTGTCTACCAACAAGAGATTTGCTTGTGAAG 906
QY 841 ATCCGTTCCAAAGAGGAGGTGTGCGAGCATCCCGGACTATACAGACATCTGCTCG 900
Db 907 ATCCGTTCCAAAGAGGAGGTGTGCGAGCATCCCGGACTATACAGACATCTGCTCG 966
QY 901 CCTCTGATGTAAAGATCTCCCAAGTTGGCAACAAGCTGTGAGATCACTGGCTTTGAAAA 960
Db 967 CCTCTGATGTAAAGATCTCCCAAGTTGGCAACAAGCTGTGAGATCACTGGCTTTGAAAA 1026
QY 961 GAGAATCTTACCGACTATCTCTATCCGAGCAGCTGAAATGACTGTGTGAAGCTGATT 1020
Db 1027 GAGAATCTTACCGACTATCTCTATCCGAGCAGCTGAAATGACTGTGTGAAGCTGATT 1086
QY 1021 TCCCAAGGGAGGTGTGAGAGCCCACTACTACAGGCTGTGAATGACACCAAAATGCTG 1080
Db 1087 TCCCAAGGGAGGTGTGAGAGCCCACTACTACAGGCTGTGAATGACACCAAAATGCTG 1146
QY 1081 TGTGCTGTCTACCCAGAGTGAAGAAACAGATTCTCTGCCAGGAGACTTCAAGGGAGCCCTC 1140
Db 1147 TGTGCTGTCTACCCAGAGTGAAGAAACAGATTCTCTGCCAGGAGACTTCAAGGGAGCCCTC 1206
QY 1141 GTCTGTTCCTTCCAAAGGCGCATGACTTTTGAATGAGTGTGAGTGGGGCGGTGATGTT 1200
Db 1207 GTCTGTTCCTTCCAAAGGCGCATGACTTTTGAATGAGTGTGAGTGGGGCGGTGATGTT 1266
QY 1201 GGCCTGAAGAGCAAGCCAGGCTGTCTACAGAGAGTCTCAACTTTTACCTGTGATCCGC 1260
Db 1267 GGCCTGAAGAGCAAGCCAGGCTGTCTACAGAGAGTCTCAACTTTTACCTGTGATCCGC 1326
QY 1261 AGTCACACCAAGAGAGATGAGCTGTGGCCCTCTGA 1296
Db 1327 AGTCACACCAAGAGAGATGAGCTGTGGCCCTCTGA 1362

US-09-480-884A-123
; Sequence 123, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121, 455C6
; CURRENT APPLICATION NUMBER: US/09/480, 884A
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-123

Query Match 100.0%; Score 1296; DB 4; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGTGGCGGCTGCTCTGTGCTGTGCTGTGAGGCACTCCAAAGGC 60
Db 67 ATGAGAGCCCTGCTGTGGCGGCTGCTCTGTGCTGTGCTGTGAGGCACTCCAAAGGC 126
QY 61 AGCAATGAACCTTCAATCAAGTCCATCGAACTGTGACTGTCTAAATGAGAAACATGTGTG 120
Db 127 AGCAATGAACCTTCAATCAAGTCCATCGAACTGTGACTGTCTAAATGAGAAACATGTGTG 186
QY 121 TCCAAAGATCTTCTCCAAATTCATCTGTGCAACTGTGCAACTGTGCAACTGTGCAACTGTG 180
Db 187 TCCAAAGATCTTCTCCAAATTCATCTGTGCAACTGTGCAACTGTGCAACTGTGCAACTGTG 246
QY 181 CACTGTGAATGATTAAGTCAAAAACCTGTGTAAGAGGAAATGTCATCTTTTACGAGGA 240
Db 247 CACTGTGAATGATTAAGTCAAAAACCTGTGTAAGAGGAAATGTCATCTTTTACGAGGA 306
QY 241 AAGGCACACCTGACACCATGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 307 AAGGCACACCTGACACCATGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
QY 301 CAGCAAAAGCTACATGCCACAGATCTGATGCTCTTCAAGCTGTGGGCTGTGGGAAACATAT 360
Db 367 CAGCAAAAGCTACATGCCACAGATCTGATGCTCTTCAAGCTGTGGGCTGTGGGAAACATAT 426
QY 361 TACTGCAAGAACCCAGACACCGGAGGGGACCTTGTGCTATGTGCAAGTGGGCTTAAAG 420
Db 427 TACTGCAAGAACCCAGACACCGGAGGGGACCTTGTGCTATGTGCAAGTGGGCTTAAAG 486
QY 421 CGCTTGTTCAGAGTGTGATGTGATGCTGCGAGATGAGAAAAAGCCCTCTCTCTCT 480
Db 487 CGCTTGTTCAGAGTGTGATGTGATGCTGCGAGATGAGAAAAAGCCCTCTCTCTCTCT 546
QY 481 CCAGAAAGATTTAAATTTTCAGTGTGCGCAAAAGACTCTGAGGCCCTGTTAAGATTAT 540
Db 547 CCAGAAAGATTTAAATTTTCAGTGTGCGCAAAAGACTCTGAGGCCCTGTTAAGATTAT 606
QY 541 GGGGGAATTCACCAACCATGAGAACAGGCCCTGTGTTGGGCACTCTACAGAGGCAC 600
Db 607 GGGGGAATTCACCAACCATGAGAACAGGCCCTGTGTTGGGCACTCTACAGAGGCAC 666
QY 601 GGGGGGGCTCTGTGACCTTACCTGTGTGTGAGAGGAGCCTCATCAGCCCTTGTGGTATC 660
Db 667 GGGGGGGCTCTGTGACCTTACCTGTGTGTGAGAGGAGCCTCATCAGCCCTTGTGGTATC 726
QY 661 AGCGGCACACAGTGTTCATGATTTACCCAAAGAAAGAGACTCATCTGTCTACCTGGT 720
Db 727 AGCGGCACACAGTGTTCATGATTTACCCAAAGAAAGAGACTCATCTGTCTACCTGGT 786

Qy	721	CGCTCAAGGCTTAATCTCCAAACCGGAAGGGAAGTAAGTTTGAGCTGGAAACCTATC	780
Db	787	CGCTCAAGGCTTAATCTCCAAACCGGAAGGGAAGTAAGTTTGAGCTGGAAACCTATC	846
Qy	781	CTACACAAGGACTACAGGCTGTGACAGCGTTGCTCACCAACGACATTCCTTGCTGAAG	840
Db	847	CTACACAAGGACTACAGGCTGTGACAGCGTTGCTCACCAACGACATTCCTTGCTGAAG	906
Qy	841	ATCGCTTCCAAAGAGGAGGAGGTGTGCGGAGCCATCCCGGACTATACAGACCATCTGCTG	900
Db	907	ATCGCTTCCAAAGAGGAGGAGGTGTGCGGAGCCATCCCGGACTATACAGACCATCTGCTG	966
Qy	901	CCCTGATGTATTAAGATCCCAAGTTTGGCACACAGCTGTGAGATCACTGGCTTTGGAAAA	960
Db	967	CCCTGATGTATTAAGATCCCAAGTTTGGCACACAGCTGTGAGATCACTGGCTTTGGAAAA	1026
Qy	961	GAGAAATTCACCGACTATCTCATTCGCGAGAGAGCTGAATAAGACTGTGTGAGACTGATT	1020
Db	1027	GAGAAATTCACCGACTATCTCATTCGCGAGAGAGCTGAATAAGACTGTGTGAGACTGATT	1086
Qy	1021	TCCCAACGGGAGTGTACAGAGCCCACTACTACGGCTCTGAAGTCAACCAAAATGCTG	1080
Db	1087	TCCCAACGGGAGTGTACAGAGCCCACTACTACGGCTCTGAAGTCAACCAAAATGCTG	1146
Qy	1081	TGTGTGTGTGACCCACAGTGTGAAAAACAGATTCTTGCCAGGAGACTCAGGGGAGACCCCTC	1140
Db	1147	TGTGTGTGTGACCCACAGTGTGAAAAACAGATTCTTGCCAGGAGACTCAGGGGAGACCCCTC	1206
Qy	1141	GTCGTGTTCCTTCGAAGCGCCGATGACTTTGACTGGAATTGTGAGCTGGGGCCGTGATGT	1200
Db	1207	GTCGTGTTCCTTCGAAGCGCCGATGACTTTGACTGGAATTGTGAGCTGGGGCCGTGATGT	1266
Qy	1201	GCCCTGAAGAGCAAGCCAGGGGCTACACGAGAGCTCAACCTTCAACCTGATCCGC	1260
Db	1267	GCCCTGAAGAGCAAGCCAGGGGCTACACGAGAGCTCAACCTTCAACCTGATCCGC	1326
Qy	1261	AGTCACACCAAGAGAGATGCGCTGGCCCTCTGA	1296
Db	1327	AGTCACACCAAGAGAGATGCGCTGGCCCTCTGA	1362

RESULT 3

US-09-542-615A-123

Sequence 123. Application US/09542615A

Patent No. 6518256

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Fan, Liqun

APPLICANT: Kalos, Michael D.

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Hosken, Nancy A.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.455C8

CURRENT APPLICATION NUMBER: US/09/542.615A

NUMBER OF SEQ ID NOS: 350

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 123

LENGTH: 2294

TYPE: DNA

ORGANISM: Homo sapien

US-09-542-615A-123

	Query Match	Similarity	Score 1296;	DB 4;	Length 2294;
	Best Local	Similarity	100.0%;	Pred. No. 0;	
	Matches 1296;	Conservative	0;	Mismatches	0; Gaps
Qy	1	ATGAGAGCCCTGCTGGCGCGCTCTGCTCTGCTGCTGAGCGACTCCAAAGC	60		
Db	67	ATGAGAGCCCTGCTGGCGCGCTCTGCTCTGCTGCTGAGCGACTCCAAAGC	126		

QY	61	GGCAATGAACTTCATCAAGTTCATCGAACTGTGACTGTCTCTAAATGGAGAAACATGTGG	120
Db	127	AGCAATGAATCTTCATCAAGTTCATCGAACTGTGACTGTCTCTAAATGGAGAAACATGTGTG	186
QY	121	TCCAACAAGTACTTCTCCAAATTCACCTGTGTGCACTGCCCCAAGAAATTCGGAGGGCAG	180
Db	187	TCCAACAAGTACTTCTCCAAATTCACCTGTGTGCACTGCCCCAAGAAATTCGGAGGGCAG	246
QY	181	CACGTGTAAATATGATTAAGTCAAAAACTGTCTATGAGGGGAATGTGCACTTTTACCGAGAGA	240
Db	247	CACGTGTAAATATGATTAAGTCAAAAACTGTCTATGAGGGGAATGTGCACTTTTACCGAGAGA	306
QY	241	AAGGCACACACTGACACCATGGGCGGGCCCTGCGCTGACCTGGAACTTGCCCACTGTCCCTT	300
Db	307	AAGGCACACACTGACACCATGGGCGGGCCCTGCGCTGACCTGGAACTTGCCCACTGTCCCTT	366
QY	301	CAGCAAAAGTACCATGCCCCACAGATCTGATGCTCTTCAGCTGGGCTGGGGAAACATATAT	360
Db	367	CAGCAAAAGTACCATGCCCCACAGATCTGATGCTCTTCAGCTGGGCTGGGGAAACATATAT	426
QY	361	TATGTCAAGAAACCCAGAGAACCGGAGGGGACCCCTGTGTCTATATGTGCAGATGTGGCCCTAAG	420
Db	427	TATGTCAAGAAACCCAGAGAACCGGAGGGGACCCCTGTGTCTATATGTGCAGATGTGGCCCTAAG	486
QY	421	CCGCTTTGTCCAAAGATGTGATGGTGCATACCTGCGCAGATGTGAAAAAAACCCCTCTCTCT	480
Db	487	CCGCTTTGTCCAAAGATGTGATGGTGCATACCTGCGCAGATGTGAAAAAAACCCCTCTCTCT	546
QY	481	CCAGAAAGATTAATAATTTCACTGTGTGGCCAAAGAATCTGTAGGCCCCGCTTTAAGTATTT	540
Db	547	CCAGAAAGATTAATAATTTCACTGTGTGGCCAAAGAATCTGTAGGCCCCGCTTTAAGTATTT	606
QY	541	GGGGGAGAATTCACCAACATGTAGAGAACCAAGCCCTGTGTGGGGCCATCTACAGAGAGGAC	600
Db	607	GGGGGAGAATTCACCAACATGTAGAGAACCAAGCCCTGTGTGGGGCCATCTACAGAGAGGAC	666
QY	601	CGGGGGGGCTCTGTCAACCTACCTAGCTGTGTGTGAGAGCAGCCTCATCAGCCCTGTGGGTGATC	660
Db	667	CGGGGGGGCTCTGTCAACCTACCTAGCTGTGTGTGAGAGCAGCCTCATCAGCCCTGTGGGTGATC	726
QY	661	AGCGCCACAACCTGCTTATGATTTATCCCAAGAAAGAGAGACTATATGTCTACTCTGGGT	720
Db	727	AGCGCCACAACCTGCTTATGATTTATCCCAAGAAAGAGAGACTATATGTCTACTCTGGGT	786
QY	721	CGCTCAAGGCTTAATCTCCACAACGCAAGGAGATGAAGTTGAGAGTGGAAAACTTAATC	780
Db	787	CGCTCAAGGCTTAATCTCCACAACGCAAGGAGATGAAGTTGAGAGTGGAAAACTTAATC	846
QY	781	CTACACANAGACTACAGGCTGTGACAGGCTTGTCTACCAACGACATTTGCTTGTCTGAAG	840
Db	847	CTACACANAGACTACAGGCTGTGACAGGCTTGTCTACCAACGACATTTGCTTGTCTGAAG	906
QY	841	ATCGGTTCCAAAGAGGGGAGGTGTGCGAGGACATCCCGAGCTTATACACACATCTGCTGTG	900
Db	907	ATCGGTTCCAAAGAGGGGAGGTGTGCGAGGACATCCCGAGCTTATACACACATCTGCTGTG	966
QY	901	CCCTGTGATGTATTAACGATCCCCAGTTTGTGGCAAGAGCTGTGATCACTGCGCTTTGGAAAA	960
Db	967	CCCTGTGATGTATTAACGATCCCCAGTTTGTGGCAAGAGCTGTGATCACTGCGCTTTGGAAAA	1026
QY	961	GAGAAATTTACACGACTATCTCTATTCGGAGCAGCTGAAATATGACTGTGTGTGAAGCTGATT	1022
Db	1027	GAGAAATTTACACGACTATCTCTATTCGGAGCAGCTGAAATATGACTGTGTGTGAAGCTGATT	1086
QY	1021	TCCCAACGGGAGTGTACAGACACCCCACTATACGGGCTGTGAAGTCAACCAACAAATGTCTG	1088
Db	1087	TCCCAACGGGAGTGTACAGACACCCCACTATACGGGCTGTGAAGTCAACCAACAAATGTCTG	1146
QY	1081	TGTGTGTGTGACCAACAGTGGAAAAACAGATTTCTGTCCAGGGAGACTCAGGGGGACCCCTC	1140
Db	1147	TGTGTGTGTGACCAACAGTGGAAAAACAGATTTCTGTCCAGGGAGACTCAGGGGGACCCCTC	1206

Accession	Sequence	Length
Db	TCACACCCGGAGTGCACACACCCCACTACAGGCTCTGAAGTCACACCAAAATGCTTG	1146
Qy	TGTGCTGTGACCCACAGTGGAAAAAGATTCTGCGCAGGGAGACTCAGGGGGACCCCTC	1140
Db	TGTGCTGTGACCCACAGTGGAAAAAGATTCTGCGCAGGGAGACTCAGGGGGACCCCTC	1206

QY 1141 GTCTGTCCTCCCAAGGCGCATGACTTTGACTGGAATTTGAGTGGGCGCCGTGATGT 1200
DB 1207 GTCTGTCCTCCCAAGGCGCGCATGACTTTGAGTGAATTTGAGTGGGCGCGTGAATGT 1266
QY 1201 GCCCTGAAGACCAAGCCGCGCTTACACGAGAGTCTTCACTTTACCTTGATCCGC 1260
DB 1267 GCCCTGAAGACCAAGCCGCGCTTACACGAGAGTCTTCACTTTACCTTGATCCGC 1326
QY 1261 AGTCACACCAAGGAAGAAATGGCTGGCCCTCTGA 1296
DB 1327 AGTCACACCAAGGAAGAAATGGCTGGCCCTCTGA 1362

RESULT 4
US-09-606-421B-123
; Sequence 123, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Ranger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-123

Query Match 100.0%; Score 1296; DB 4; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGGCGCGCTGCTTCTGCGTCCGTGAGCACTCCAAAGGC 60
DB 67 ATGAGAGCCCTGCTGGCGCGCTGCTTCTGCGTCCGTGAGCACTCCAAAGGC 126
QY 61 AGCAATGAAGTTCATCAAGTTCATGGAATCTGTAATGAGAGAACTGTGTG 120
DB 127 AGCAATGAAGTTCATCAAGTTCATGGAATCTGTAATGAGAGAACTGTGTG 186
QY 121 TCCAAACAAGTCTTCTCAACATTTCACTGGTGAACGCGCAAGAAATTCGAGAGCGAG 180
DB 187 TCCAAACAAGTCTTCTCAACATTTCACTGGTGAACGCGCAAGAAATTCGAGAGCGAG 246
QY 181 CACTGTGAATTAATAAGTCAAAAACCTGTATGAGGGGAATGCTCACTTTTACGAGGA 240
DB 247 CACTGTGAATTAATAAGTCAAAAACCTGTATGAGGGGAATGCTCACTTTTACGAGGA 306
QY 241 AAGGCCAGCATGACCAACCAATGGGCGGCGCTGCTGCGCTTGAACCTCTGCACTGCTT 300
DB 307 AAGGCCAGCATGACCAACCAATGGGCGGCGCTGCTGCGCTTGAACCTCTGCACTGCTT 366
QY 301 CAGCAAAAGTACCATGCCCAAGATCTGATCTTCACTGGCGCTGAGGAAACATATAT 360
DB 367 CAGCAAAAGTACCATGCCCAAGATCTGATCTTCACTGGCGCTGAGGAAACATATAT 426
QY 361 TACTGAGGAACCCCAAGAACCGGAGCGAAGCTGTGCTATGTGAGGTGGGCTTAAG 420
DB 427 TACTGAGGAACCCCAAGAACCGGAGCGAAGCTGTGCTATGTGAGGTGGGCTTAAG 486
QY 421 CCGCTTGTCCAAGAGTGCATGTGATGACTGCGGAGATGAGAAAAAGCCCTCTCTCT 480

DB 487 CCGCTTGTCCAAGAGTGCATGTGATGACTGCGGAGATGAGAAAAAGCCCTCTCTCT 546
QY 481 CCAGAAAGATTTAAATTTAGTGTGCGCAAAAAGATCTGAGGGCGCGCTTTAAGTTAT 540
DB 547 CCAGAAAGATTTAAATTTAGTGTGCGCAAAAAGATCTGAGGGCGCGCTTTAAGTTAT 606
QY 541 GGGGGAAATTTACCAACATCGAGAACCAAGCCCTGTTGGCGGCATCTACAGAGGAGC 600
DB 607 GGGGGAAATTTACCAACATCGAGAACCAAGCCCTGTTGGCGGCATCTACAGAGGAGC 666
QY 601 CGGGGGGGCTGTGATCACTGATGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
DB 667 CGGGGGGGCTGTGATCACTGATGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 726
QY 661 AGCGCCACACATGCTTATTTATTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 727 AGCGCCACACATGCTTATTTATTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 786
QY 721 CGCTCAAGGCTTAACTCCAAACGCAAGGAGAGATGAAGTTGAGTGAAGAAACCTAATC 780
DB 787 CGCTCAAGGCTTAACTCCAAACGCAAGGAGAGATGAAGTTGAGTGAAGAAACCTAATC 846
QY 781 CTACACAAAGACTACAGGCTGACACGCTTGTCTACCAACAGCAATTTGCTTGTAG 840
DB 847 CTACACAAAGACTACAGGCTGACACGCTTGTCTACCAACAGCAATTTGCTTGTAG 906
QY 841 ATCCGTTCCAAAGAGAGGAGGTGTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 907 ATCCGTTCCAAAGAGAGGAGGTGTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
QY 901 CCCTGAGTAAATTAAGATTTCCCAAGTGTGAGCAAGGCTGTGATCACTGGCTTGAAGA 960
DB 967 CCCTGAGTAAATTAAGATTTCCCAAGTGTGAGCAAGGCTGTGATCACTGGCTTGAAGA 1026
QY 961 GAGAAATTTACCGACTATCTCTATCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
DB 1027 GAGAAATTTACCGACTATCTCTATCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1086
QY 1021 TCCCAAGGAGGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1087 TCCCAAGGAGGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1146
QY 1081 TGTGTGTGATGACCAAGTGAAGAAAGATTTCTGCGCAAGGAGGAGGAGGAGGAGGAG 1140
DB 1147 TGTGTGTGATGACCAAGTGAAGAAAGATTTCTGCGCAAGGAGGAGGAGGAGGAGGAG 1206
QY 1141 GTCTGTTCCCTCCAAAGCCGCGATGACTTTGACTGGAATTTGAGCTGGGCGCGTGAAT 1200
DB 1207 GTCTGTTCCCTCCAAAGCCGCGATGACTTTGACTGGAATTTGAGCTGGGCGCGTGAAT 1266
QY 1201 GCCCTGAAGACCAAGCCGCGCTTACACGAGAGTCTCACTTTTACCTTGATCCGC 1260
DB 1267 GCCCTGAAGACCAAGCCGCGCTTACACGAGAGTCTCACTTTTACCTTGATCCGC 1326
QY 1261 AGTCACACCAAGGAAGAAATGGCTGGCCCTCTGA 1296
DB 1327 AGTCACACCAAGGAAGAAATGGCTGGCCCTCTGA 1362

RESULT 5
US-09-023-655-1217
; Sequence 1217, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Selhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

[illegible]

QY 241 AA

Query	Db	Query Match	Best Local Similarity	Matches 1294	Conservative	Score 1292.8	DB 4	Length 1475	Indels	Gaps	0
1	ATGAGAGCCCTCTGGCGCGCCTGCTCTCTGTCGTCGTCGTGAGGAGCTCCAAAGCC	99.8%	99.8%	0	0	2	0	0	0	0	0
81	ATGAGAGCCCTCTGGCGCGCCTGCTCTCTGTCGTCGTCGTGAGGAGCTCCAAAGCC	99.8%	99.8%	0	0	2	0	0	0	0	0
61	AGCAATGAACTTCATCAAGTTCCATCGAATCTGTGACTGTCTTAATGAGGAACATGTGTG	99.8%	99.8%	0	0	2	0	0	0	0	0
141	AGCAATGAACTTCATCAAGTTCCATCGAATCTGTGACTGTCTTAATGAGGAACATGTGTG	99.8%	99.8%	0	0	2	0	0	0	0	0
121	TCCAAACAGTACTTCTCCAAACATTCACGTGTGCAACATCCCAAGAAATTCGGAAGGCAG	99.8%	99.8%	0	0	2	0	0	0	0	0
201	TCCAAACAGTACTTCTCCAAACATTCACGTGTGCAACATCCCAAGAAATTCGGAAGGCAG	99.8%	99.8%	0	0	2	0	0	0	0	0
181	CACGTGTGAATAGATAGTCAAAAACCTGTATGAGGGAATGTCACTTTTACCGAGGA	99.8%	99.8%	0	0	2	0	0	0	0	0
261	CACGTGTGAATAGATAGTCAAAAACCTGTATGAGGGAATGTCACTTTTACCGAGGA	99.8%	99.8%	0	0	2	0	0	0	0	0
241	AAGGCAACACTGACACCATGGGCGGCGCTGCTGCTGCTGGAACCTTGGCACTGTCTT	99.8%	99.8%	0	0	2	0	0	0	0	0
321	AAGGCAACACTGACACCATGGGCGGCGCTGCTGCTGCTGGAACCTTGGCACTGTCTT	99.8%	99.8%	0	0	2	0	0	0	0	0

QY 301 CAGCAACGACGATCCCAAGATCTGATCTTCACTGGGCTGGGAAACATTAAT 360
 DB 381 CAGCAACGACGATCCCAAGATCTGATCTTCACTGGGCTGGGAAACATTAAT 440
 QY 361 TACTGAGGAAACCAAGCAACCGGAGGCGACCTGTGTCTATGTGAGGTGGCTTAAG 420
 DB 441 TACTGAGGAAACCAAGCAACCGGAGGCGACCTGTGTCTATGTGAGGTGGCTTAAG 500
 QY 421 CCGCTGTCCAGAGTGATGTGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 501 CCGCTGTCCAGAGTGATGTGATGATGATGATGATGATGATGATGATGATGATGAT 560
 QY 481 CCGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
 DB 561 CCGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 620
 QY 541 GGGGAGAAATTCACCAACGATCGAAGAACGAGCCCTGGTTTGGGCACTTACAGAGGAC 600
 DB 621 GGGGAGAAATTCACCAACGATCGAAGAACGAGCCCTGGTTTGGGCACTTACAGAGGAC 680
 QY 601 CCGGGGGGGCTGTGTCACTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 DB 681 CCGGGGGGGCTGTGTCACTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 740
 QY 661 AGGCGCACACACTGCTTCACTTATTAATCCAAAGAGAGACTTACATCTGTCTAGGT 720
 DB 741 AGGCGCACACACTGCTTCACTTATTAATCCAAAGAGAGACTTACATCTGTCTAGGT 800
 QY 721 CGCTCAAGGCTTAATCTCAACGACGAGGGGAGATGATGATGATGATGATGATGATGAT 780
 DB 801 CGCTCAAGGCTTAATCTCAACGACGAGGGGAGATGATGATGATGATGATGATGATGAT 860
 QY 781 CTACACAGAGACTACAGCGCTGACACGCTTGTCTACACACAGACTTGTCTGTGAAG 840
 DB 861 CTACACAGAGACTACAGCGCTGACACGCTTGTCTACACACAGACTTGTCTGTGAAG 920
 QY 841 ATCCGCTTCCAGAGAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
 DB 921 ATCCGCTTCCAGAGAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 980
 QY 901 CCCTCATGTATACGATCCCAAGTTTGGCACAAGCTGTGATGATGATGATGATGATGAT 960
 DB 981 CCCTCATGTATACGATCCCAAGTTTGGCACAAGCTGTGATGATGATGATGATGATGAT 1040
 QY 961 GAGAAATTCACGACATCTCTATTCGAGAGACGCTGAAAATGACTGTGTGAAGCTGATT 1020
 DB 1041 GAGAAATTCACGACATCTCTATTCGAGAGACGCTGAAAATGACTGTGTGAAGCTGATT 1100
 QY 1021 TCCCAACCGGAGAGTGTGACGACGCCCACTACTACTGTGTGAAGTCAACCAAAATGCTG 1080
 DB 1101 TCCCAACCGGAGAGTGTGACGACGCCCACTACTACTGTGTGAAGTCAACCAAAATGCTG 1160
 QY 1081 TGT 1140
 DB 1161 TGT 1220
 QY 1141 GTCTGT 1200
 DB 1221 GTCTGT 1280
 QY 1201 GCGCTGAGAGCAAGCGAGCGCTGTACAGAGAGTCTCACACTTCTTAACCTTGAATCCG 1260
 DB 1281 GCGCTGAGAGCAAGCGAGCGCTGTACAGAGAGTCTCACACTTCTTAACCTTGAATCCG 1340
 QY 1261 AGTCAACCAAG 1296
 DB 1341 AGTCAACCAAG 1376

RESULT 8
 US-09-480-884A-122
 ; Sequence 122, Application US/09480884A

; Patent No. 6482597
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Koles, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; FILE REFERENCE: 210121.455C6
 ; CURRENT APPLICATION NUMBER: US/09/480,884A
 ; CURRENT FILING DATE: 2001-08-27
 ; NUMBER OF SEQ ID NOS: 330
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 122
 ; LENGTH: 1475
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-480-884A-122

Query Match 99.8%; Score 1292.8; DB 4; Length 1475;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGGCGGGCTGTCTCTGTGCTCTGTGTGTGTGTGTGTGTGTGTGTGT 60
 DB 81 ATGAGAGCCCTGCTGGCGGGCTGTCTCTGTGCTCTGTGTGTGTGTGTGTGTGTGTGT 140
 QY 61 AGCAATGAATCTTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTCAT 120
 DB 141 AGCAATGAATCTTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTCAT 200
 QY 121 TCCACAAAGTACTTCTTCCAAATTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
 DB 201 TCCACAAAGTACTTCTTCCAAATTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 260
 QY 181 CACTGTGAATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 240
 DB 261 CACTGTGAATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 320
 QY 241 AAGGCGAGACTGTGACACATGAGGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
 DB 321 AAGGCGAGACTGTGACACATGAGGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 380
 QY 301 CAGCAACGATACATCCCAAGATCTGATGTCTTCACTGTGTGTGTGTGTGTGTGTGTGTGT 360
 DB 381 CAGCAACGATACATCCCAAGATCTGATGTCTTCACTGTGTGTGTGTGTGTGTGTGTGTGT 440
 QY 361 TACTGAGGAAACCAAGCAACCGGAGGCGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
 DB 441 TACTGAGGAAACCAAGCAACCGGAGGCGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 500
 QY 421 CCGCTGTCCAGAGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 501 CCGCTGTCCAGAGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 560
 QY 481 CCGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
 DB 561 CCGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 620
 QY 541 GGGGAGAAATTCACCAACGATCGAAGAACGAGCCCTGGTTTGGGCACTTACAGAGGAC 600
 DB 621 GGGGAGAAATTCACCAACGATCGAAGAACGAGCCCTGGTTTGGGCACTTACAGAGGAC 680
 QY 601 CCGGGGGGGCTGTGTCACTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 DB 681 CCGGGGGGGCTGTGTCACTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 740
 QY 661 AGGCGCACACACTGCTTCACTTATTAATCCAAAGAGAGACTTACATCTGTCTAGGT 720
 DB 741 AGGCGCACACACTGCTTCACTTATTAATCCAAAGAGAGACTTACATCTGTCTAGGT 800
 QY 721 CGCTCAAGGCTTAATCTCAACGACGAGGGGAGATGATGATGATGATGATGATGATGATGAT 780

Db 801 CGCTAAGGCTTAACCTCAACGCAAGGAGAGATGTTGAGTGGAAAACTCATC 860
Qy 781 CTACCAAGAAGCTACAGCGCTGACACGCTTCTCTACCAACAAGCATTTGCTGTAG 840
Db 861 CTACCAAGAAGCTACAGCGCTGACACGCTTCTCTACCAACAAGCATTTGCTGTAG 920
Qy 841 ATCCGTTCCAAAGGAGGAGGTGTGCGGACCATTCGCGACTTATACAGACCATTCGCTG 900
Db 921 ATCCGTTCCAAAGGAGGAGGTGTGCGGACCATTCGCGACTTATACAGACCATTCGCTG 980
Qy 901 CCTCGATGTATTAACGATCCCAAGTTTGGCAACAAGTGTGATCACTGGCTTTGAAAA 960
Db 981 CCTCGATGTATTAACGATCCCAAGTTTGGCAACAAGTGTGATCACTGGCTTTGAAAA 1040
Qy 961 GAGAAATTTACCGACTATCTTATTCGAGAGCACTGAAAAATGACTGTTGAAAGCTGAT 1020
Db 1041 GAGAAATTTACCGACTATCTTATTCGAGAGCACTGAAAAATGACTGTTGAAAGCTGAT 1100
Qy 1021 TCCCAACGGAGAGTGTACAGACCCCACTACAGGCTGTAAAGTCAACCAAAATGCTG 1080
Db 1101 TCCCAACGGAGAGTGTACAGACCCCACTACAGGCTGTAAAGTCAACCAAAATGCTG 1160
Qy 1081 TGTCTGTCTGACCCCAAGTGAACAAAGATTCTGCGAGGAGAGCTCAGGGGACCCCTC 1140
Db 1161 TGTCTGTCTGACCCCAAGTGAACAAAGATTCTGCGAGGAGAGCTCAGGGGACCCCTC 1220
Qy 1141 GTCTGTTCCCTCCAAAGGCGCGATGACTTTGAGCTGAGTGTGAGTGTGAGTGTGAGTGT 1200
Db 1221 GTCTGTTCCCTCCAAAGGCGCGATGACTTTGAGCTGAGTGTGAGTGTGAGTGTGAGTGT 1280
Qy 1201 GCCCTGAAGGACAAAGCCAGGCGTCTACACGAGAGTCTACACTTCTTACCTGAGATCCG 1260
Db 1281 GCCCTGAAGGACAAAGCCAGGCGTCTACACGAGAGTCTACACTTCTTACCTGAGATCCG 1340
Qy 1261 AGTACACCAAGGAAGAAATGGCTGCGCTCTGA 1296
Db 1341 AGTACACCAAGGAAGAAATGGCTGCGCTCTGA 1376

RESULT 9
US-09-542-615A-122
; Sequence 122, Application US/09542615A
; Patent No. 6518236
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-542-615A-122

Query Match 99.8%; Score 1292.8; DB 4; Length 1475;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGAGCCCTGTGCGCGGCTGCTTCTGCTGCTGTGTGAGGACTCCAAAGGC 60
Db 81 ATGAGAGCCCTGTGCGCGGCTGCTTCTGCTGCTGTGTGAGGACTCCAAAGGC 140
Qy 61 AGCAATGAATTCATCAAGTTCATGCACTGTGACTGTCTAAATGAGAGAAACATGTGTG 120

Db 141 AGCAATGAATTCATCAAGTTCATGCACTGTGACTGTCTAAATGAGAGAAACATGTGTG 200
Qy 121 TCCAAATGATCTTCTCCAAACATTCACGTGTGCAACTGCCCAAGAAATTCGAGGCGAG 180
Db 201 TCCAAATGATCTTCTCCAAACATTCACGTGTGCAACTGCCCAAGAAATTCGAGGCGAG 260
Qy 181 CACTGTGAATATGATTAAGTCAAAAACTGCTATGAGGGAATGTCACTTTTACCGAGGA 240
Db 261 CACTGTGAATATGATTAAGTCAAAAACTGCTATGAGGGAATGTCACTTTTACCGAGGA 320
Qy 241 AAGGCGACGACTGACACCAATGGCGCGGCTGTGCTGCTGCTGCAACTGCTGCTT 300
Db 321 AAGGCGACGACTGACACCAATGGCGCGGCTGTGCTGCTGCTGCAACTGCTGCTT 380
Qy 301 CAGCAAACTACCATGACCAAGATCTGATGCTCTTCAAGTGGGCTGAGGGAACATAT 360
Db 381 CAGCAAACTACCATGACCAAGATCTGATGCTCTTCAAGTGGGCTGAGGGAACATAT 440
Qy 361 TACTGAGGAACCCAGACAAACCGAGGAGCCCTGCTGTCTATGTGCAAGTGGGCTTAAAG 420
Db 441 TACTGAGGAACCCAGACAAACCGAGGAGCCCTGCTGTCTATGTGCAAGTGGGCTTAAAG 500
Qy 421 CGCTGTGTCAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Db 501 CGCTGTGTCAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 560
Qy 481 CCAGAAATTAATAATTTCAAGTGTGCGCAAAAGACTGAGGCCCCGCTTTAAGATTAT 540
Db 561 CCAGAAATTAATAATTTCAAGTGTGCGCAAAAGACTGAGGCCCCGCTTTAAGATTAT 620
Qy 541 GGGGGAATTAATCAACCATGAGAACCAAGCCCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 600
Db 621 GGGGGAATTAATCAACCATGAGAACCAAGCCCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 680
Qy 601 CGGGGGGCTGTGACCTACGTGTGTGAGAGGAGCTCATCAGCCCTTGTGCTGATG 660
Db 681 CGGGGGGCTGTGACCTACGTGTGTGAGAGGAGCTCATCAGCCCTTGTGCTGATG 740
Qy 661 AGGCGCACACTGTGCTTATGATTAATCCCAAGAAAGAGAGCTACATGCTTACCTGGGT 720
Db 741 AGGCGCACACTGTGCTTATGATTAATCCCAAGAAAGAGAGCTACATGCTTACCTGGGT 800
Qy 721 CGCTCAAGGCTTAATCTCAACAGCAAGGAGAGTGAAGTTGAGTGTGAAAACTTAATC 780
Db 801 CGCTCAAGGCTTAATCTCAACAGCAAGGAGAGTGAAGTTGAGTGTGAAAACTTAATC 860
Qy 781 CTACCAAGAAGCTACAGCGCTGACACGCTTCTCTACCAACAAGCATTTGCTGTAG 840
Db 861 CTACCAAGAAGCTACAGCGCTGACACGCTTCTCTACCAACAAGCATTTGCTGTAG 920
Qy 841 ATCCGTTCCAAAGGAGGAGGTGTGCGGACCATTCGCGACTTATACAGACCATTCGCTG 900
Db 921 ATCCGTTCCAAAGGAGGAGGTGTGCGGACCATTCGCGACTTATACAGACCATTCGCTG 980
Qy 901 CCTCGATGTATTAACGATCCCAAGTTTGGCAACAAGTGTGATCACTGGCTTTGAAAA 960
Db 981 CCTCGATGTATTAACGATCCCAAGTTTGGCAACAAGTGTGATCACTGGCTTTGAAAA 1040
Qy 961 GAGAAATTTACCGACTATCTTATTCGAGAGCACTGAAAAATGACTGTTGAAAGCTGAT 1020
Db 1041 GAGAAATTTACCGACTATCTTATTCGAGAGCACTGAAAAATGACTGTTGAAAGCTGAT 1100
Qy 1021 TCCCAACGGAGAGTGTACAGACCCCACTACAGGCTGTAAAGTCAACCAAAATGCTG 1080
Db 1101 TCCCAACGGAGAGTGTACAGACCCCACTACAGGCTGTAAAGTCAACCAAAATGCTG 1160
Qy 1081 TGTCTGTCTGACCCCAAGTGAACAAAGATTCTGCGAGGAGAGCTCAGGGGACCCCTC 1140
Db 1161 TGTCTGTCTGACCCCAAGTGAACAAAGATTCTGCGAGGAGAGCTCAGGGGACCCCTC 1220
Qy 1141 GTCTGTTCCCTCCAAAGGCGCGATGACTTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1200

Db 1221 GTCGTGTCCTCCAGGCGCCGATGATCTTGATGGAATGAGCTGGGGCCGTGATGT 1280
Qy 1201 GGCCTAAGGACAGAGCGGCTCTACAGAGTCTCACTTCTTACCTTGATGATCCG 1260
Db 1281 GGCCTGAAGACAGAGCGGCTCTACAGAGTCTCACTTCTTACCTTGATGATCCG 1340
Qy 1261 AGTCACACCAAGAGAGATGAGCTGGCCCTCTGA 1296
Db 1341 AGTCACACCAAGAGAGATGAGCTGGCCCTCTGA 1376
RESULT 10
US-09-606-421B-122
; Sequence 122, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606.421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-122
Query Match 99.8%; Score 1292.8; DB 4; Length 1475;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAGAGCCCTGTGGCGGCTGTCTCTGTGCTGCTGGTGGAGCGATCCAAAGG 60
Db 81 ATGAGAGCCCTGTGGCGGCTGTCTCTGTGCTGCTGGTGGAGCGATCCAAAGG 140
Qy 61 AGCAATGAATCTTCATCAAGTTCATCGAATCTGATCTGTCTTAATGAGAGACATGTGT 120
Db 141 AGCAATGAATCTTCATCAAGTTCATCGAATCTGATCTGTCTTAATGAGAGACATGTGT 200
Qy 121 TCCAAAGATGATCTTCTTCAACATTCACCTGTCACCTGCCAAAGAAATTCGAGGCGAG 180
Db 201 TCCAAAGATGATCTTCTTCAACATTCACCTGTCACCTGCCAAAGAAATTCGAGGCGAG 260
Qy 181 CACTGTGAATAGATTAAGTCAAAAACCTGTATAGAGGGAAATGTAATTTTACCGAGA 240
Db 261 CACTGTGAATAGATTAAGTCAAAAACCTGTATAGAGGGAAATGTAATTTTACCGAGA 320
Qy 241 AAGGCGAGCACTGACACATGAGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 321 AAGGCGAGCACTGACACATGAGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380
Qy 301 CAGCAAGATGATCTTCTTCAACATTCATGATCTTCTTCAAGTGGGCTGGGAAACATTAAT 360
Db 381 CAGCAAGATGATCTTCTTCAACATTCATGATCTTCTTCAAGTGGGCTGGGAAACATTAAT 440
Qy 361 TACTGAGGAAACCAAGACAGAGCGAGCGACCTGTGATGATGAGGAGGAGGAGGAGGAG 420
Db 441 TACTGAGGAAACCAAGACAGAGCGAGCGACCTGTGATGATGAGGAGGAGGAGGAGGAG 500
Qy 421 CGCTTGTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 501 CGCTTGTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 560

Qy 481 CCAGAAATTAATTAATTTAGTGTGGCAAAAAGACTGTAGGGCCCGCTTTAATTAATT 540
Db 561 CAGAAAGATTAATTAATTTAGTGTGGCAAAAAGACTGTAGGGCCCGCTTTAATTAATT 620
Qy 541 GGGGAGAAATTCACCAACATGAGAAACAGCCCTGTGTTGGGCAATCTACAGAGGAC 600
Db 621 GGGGAGAAATTCACCAACATGAGAAACAGCCCTGTGTTGGGCAATCTACAGAGGAC 680
Qy 601 CGGGGGGGCTGTGTACCTTACCTTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 681 CGGGGGGGCTGTGTACCTTACCTTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 740
Qy 661 AGGCGCACACACGCTTCAATGATTAATCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAG 720
Db 741 AGGCGCACACACGCTTCAATGATTAATCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAG 800
Qy 721 CGCTCAAGGCTTAATCTCAACACGCAAGGAGAGATGAATTTGAGGTGAGAAACCTAATC 780
Db 801 CGCTCAAGGCTTAATCTCAACACGCAAGGAGAGATGAATTTGAGGTGAGAAACCTAATC 860
Qy 781 CTACACAGAGACTACAGCGCTGACAGCTTGTCTACACACAGACATTTGCTTGTGAAG 840
Db 861 CTACACAGAGACTACAGCGCTGACAGCTTGTCTACACACAGACATTTGCTTGTGAAG 920
Qy 841 ATCCGTTCCAGAGAGGAGGAGTGTGGCAGCCATCCCGGACTTACAGACATCTGCTG 900
Db 921 ATCCGTTCCAGAGAGGAGGAGTGTGGCAGCCATCCCGGACTTACAGACATCTGCTG 980
Qy 901 CCTCGATGATTAACGATCCCGAGTTGGCAACAGCTGTGAGATCACTGGCTTTGAGAAA 960
Db 981 CCTCGATGATTAACGATCCCGAGTTGGCAACAGCTGTGAGATCACTGGCTTTGAGAAA 1040
Qy 961 GAGAAATTAACGATCTATCTTATCCGAGAGCAGCTGAAAATGACTGTGTGAAGTGAAT 1020
Db 1041 GAGAAATTAACGATCTATCTTATCCGAGAGCAGCTGAAAATGACTGTGTGAAGTGAAT 1100
Qy 1021 TCCACCGGAGAGTGTAGAGAGCCCACTACTAGGCTGTGAATGACACCAAAATGCTG 1080
Db 1101 TCCACCGGAGAGTGTAGAGAGCCCACTACTAGGCTGTGAATGACACCAAAATGCTG 1160
Qy 1081 TGTGCTGTGACCCACAGTGAAGAAACAGATTCTGTGCAGGAGAGCTCAGGGGAGACCC 1140
Db 1161 TGTGCTGTGACCCACAGTGAAGAAACAGATTCTGTGCAGGAGAGCTCAGGGGAGACCC 1220
Qy 1221 GTCGTGTCCTCCAAAGCGGCAATGACTTGAATGAGTGTGAGCTGGGGCGGTGATGT 1280
Qy 1201 GGCCTGAAGACAGAGCGGCTTACAGAGATCTACACTTCTTACCTTGATGATCCG 1260
Db 1281 GGCCTGAAGACAGAGCGGCTTACAGAGATCTACACTTCTTACCTTGATGATCCG 1340
Qy 1261 AGTCACACCAAGAGAGATGAGGCTGGCCCTCTGA 1296
Db 1341 AGTCACACCAAGAGAGATGAGGCTGGCCCTCTGA 1376
RESULT 11
US-09-221-107-122
; Sequence 122, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C2
; CURRENT APPLICATION NUMBER: US/09/221.107
; CURRENT FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA

Db	547	CCAGAGAAATTAAATTTTCAGTGTGGCCAAAAGACTCTGAGGCCCGCTTTAAGATTATTT	606
Qy	541	GGGGGAGAAATTCACCAACATCGAGAACAGACCCTGGTTTTGCCGCCATCTAACAGAGACAC	600
Db	607	GGGGGAGAAATTCACCAACATCGAGAACAGACCCTGGTTTTGCCGCCATCTAACAGAGACAC	666
Qy	601	CGGGGGGGGCTCTGTCACTACGTGTGTGGAGGAGAGCTTATCAGCCCTTGTGGGTGATC	660
Db	667	CGGGGGGGGCTCTGTCACTACGTGTGTGGAGGAGAGCTTATCAGCCCTTGTGGGTGATC	726
Qy	661	AGCGGCACACATGCTCTCATGTATTACCCAAAGAAAGAGAGACTACATGCTTACCTGGGT	720
Db	727	AGCGGCACACATGCTCTCATGTATTACCCAAAGAAAGAGAGACTACATGCTTACCTGGGT	786
Qy	721	CGCTCAAGGCTTTACTCTCCAAACGCAAGGAGAGATGAAATTTAGGTGGAAAACTTAATC	780
Db	787	CGCTCAAGGCTTTACTCTCCAAACGCAAGGAGAGATGAAATTTAGGTGGAAAACTTAATC	846
Qy	781	CTACACAAAGACTTACAGCGCTGTACACAGCTTGTCTCACACACACGACTTGCCTGTGAAG	840
Db	847	CTACACAAAGACTTACAGCGCTGTACACAGCTTGTCTCACACAAATGACATTGCTTGTGAAG	906
Qy	841	ATCGCTTCCAAAGAGGAGGCGGTGTGCGGAGCCATCCCGGACTATACAGACATCTGCGCTG	900
Db	907	ATCGCTTCCAAAGAGGAGGCGGTGTGCGGAGCCATCCCGGACTATACAGACATCTGCGCTG	966
Qy	901	CCCTGATGTATTAACGATCCCCAGTTTGGACAAGCTGTGAATCACTGCTTTGGAAAA	960
Db	967	CCCTGATGTATTAACGATCCCCAGTTTGGACAAGCTGTGAATCACTGCTTTGGAAAA	1022
Qy	961	GAGAAATTTACCCGACTATCTCTATCCGAGCAGCTGAAATGACTGTTGTGAAGCTGATT	1022
Db	1027	GAGAAATTTACCCGACTATCTCTATCCGAGCAGCTGAAATGACTGTTGTGAAGCTGATT	1088
Qy	1021	TCCCAACCGGAGTGTTCAGCAGGCCCACTATACGGCTCTGAAGTCAACCAACMAATGCTG	1088
Db	1087	TCCCAACCGGAGTGTTCAGCAGGCCCACTATACGGCTCTGAAGTCAACCAACMAATGCTG	1144
Qy	1081	TGTGTGTGTGACCCCAAGTGGAAAAAGATTCTCTGACAGGAGAGACTCAGGGGAGACCCCTC	1144
Db	1147	TGTGTGTGTGACCCCAAGTGGAAAAAGATTCTCTGACAGGAGAGACTCAGGGGAGACCCCTC	1200
Qy	1141	GTCGTGTTCCCTCCAAAGGCGCGACTGACTTGACTGGAATTGTGAGCTGAGGCGCTGGATGT	1200
Db	1207	GTCGTGTTCCCTCCAAAGGCGCGACTGACTTGACTGGAATTGTGAGCTGAGGCGCTGGATGT	1266
Qy	1201	GCCCTGAAGAGACACCGCAGCGCTTACACGAGAGTCTCACTTCTTACCTTGAATCCGC	1266
Db	1267	GCCCTGAAGAGACACCGCAGCGCTTACACGAGAGTCTCACTTCTTACCTTGAATCCGC	1322
Qy	1261	AGTACACCAAGAGAGAAATGGCCCTGGCCCTCTGA	1296
Db	1327	AGTACACCAAGAGAGAAATGGCCCTGGCCCTCTGA	1362

RESULT 13
 5219569-1
 Patent No. 5219569
 APPLICANT: BLABER, MICHAEL, HEYNEKER, HERBERT L.; VEHAR,
 GORDON A.
 TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
 NUMBER OF SEQUENCES: 6
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/766,858
 FILING DATE: 16-AUG-1985
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 725,468
 FILING DATE: 22-APR-1985
 SEQ ID NO.:1
 LENGTH: 1372
 5219569-1

Query Match	99.4%	Score 1288	DB 6	Length 1372	
Best Local Similarity	99.6%	Pred. No. 0			
Matches 1291	Conservative	0	Mismatches	5	Indels
					Gaps
					0
QY	1	ATGAGAGCCCTGCTGGCGCGCTTCTCTGCGCGTCTGATCGTAGCGACTCCAAAGGC	60		
DB	77	ATGAGAGCCCTGCTGGCGCGCTTCTCTGCGCTGATCGTAGCGACTCCAAAGGC	136		
QY	61	AGCAATGAATTCATCAAGTTTCATCGAATGTGACTGTCTTAATGAGGAACATGTGTG	120		
DB	137	AGCAATGAATTCATCAAGTTTCATCGAATGTGACTGTCTTAATGAGGAACATGTGTG	196		
QY	121	TCCAACAAGTACTTCTCCAACTTACTGTGTGCACTGCCCAAGAAATTGGAGGCGAG	180		
DB	197	TCCAACAAGTACTTCTCCAACTTACTGTGTGCACTGCCCAAGAAATTGGAGGCGAG	256		
QY	181	CAGTGTGAATATGATTAAGTCAAAAACTGTATGAGGGGAAATGTGCACTTTTACGAGGA	240		
DB	257	CAGTGTGAATATGATTAAGTCAAAAACTGTATGAGGGGAAATGTGCACTTTTACGAGGA	316		
QY	241	AAGGCGACACTGACACCATGAGGCCGCGCTGCTGCCCTGGAACCTGTCCTT	300		
DB	317	AAGGCGACACTGACACCATGAGGCCGCGCTGCTGCCCTGGAACCTGTCCTT	376		
QY	301	CAGCAAAAGTACATGATGCCCAAGATCTATGCTCTTACGCTGGGCTTGGGAAACATAT	360		
DB	377	CAGCAAAAGTACATGATGCCCAAGATCTATGCTCTTACGCTGGGCTTGGGAAACATAT	436		
QY	361	TACTGCAGAAACCCAGACAAACCGGAGGGGACCCCTGTGTCTATGTGACAGTGGGCTTAAG	420		
DB	437	TACTGCAGAAACCCAGACAAACCGGAGGGGACCCCTGTGTCTATGTGACAGTGGGCTTAAG	496		
QY	421	CGCGTTGTCCAAGAGTGCATGTGCATGACTGCGCAGATGGAATAAAACCCCTCTCTCT	480		
DB	497	CGCGTTGTCCAAGAGTGCATGTGCATGACTGCGCAGATGGAATAAAACCCCTCTCTCT	556		
QY	481	CCAGAGAATTTAAATTTCACTGTGTGCCAAAGAAGCTCTGAGGCCCGCTTTAAGATTAT	540		
DB	557	CCAGAGAATTTAAATTTCACTGTGTGCCAAAGAAGCTCTGAGGCCCGCTTTAAGATTAT	616		
QY	541	GGGGGAGAAATCACACCATGAGAAACGAGCCGTGTGTGGGGCATCTACAGAGGAGAC	600		
DB	617	GGGGGAGAAATCACACCATGAGAAACGAGCCGTGTGTGGGGCATCTACAGAGGAGAC	676		
QY	601	CGGGGGGCTCTGTCAACCTACGTGTGTGGAAGAGAGCTTACAGCCCTTGTGGGTGATC	660		
DB	677	CGGGGGGCTCTGTCAACCTACGTGTGTGGAAGAGAGCTTACAGCCCTTGTGGGTGATC	736		
QY	661	AGCGGCACACCTGCTTCAATTGATTATCCCAAGAAGAGAGCTACATGCTTACTGTGGGT	720		
DB	737	AGCGGCACACCTGCTTCAATTGATTATCCCAAGAAGAGAGCTACATGCTTACTGTGGGT	796		
QY	721	CGCTCAAGGCTTAACTCCAAACGCAAGGAGAGATGAAGTTGAGGTGGAATCCTATC	780		
DB	797	CGCTCAAGGCTTAACTCCAAACGCAAGGAGAGATGAAGTTGAGGTGGAATCCTATC	856		
QY	781	CTACACAAGAGCTACAGGCTGACAGCGTTTCTCACCAACAAGCATTTGCTTGGTGAAG	840		
DB	857	CTACACAAGAGCTACAGGCTGACAGCGTTTCTCACCAACAAGCATTTGCTTGGTGAAG	916		
QY	841	ATCCGTTTCAGAGAGGAGGTGTGCGAGCCATCCCGGACTATACAGACCATGTGCTG	900		
DB	917	ATCCGTTTCAGAGAGGAGGTGTGCGAGCCATCCCGGACTATACAGACCATGTGCTG	976		
QY	901	CCCTCGATGTATTAACGATCCCAAGTTTGGCAACAAGCTGTGAGATCACGTGGCTTTGGA	960		
DB	977	CCCTCGATGTATTAACGATCCCAAGTTTGGCAACAAGCTGTGAGATCACGTGGCTTTGGA	1036		
QY	961	GAGAAATTCACGACTATCTCTATTCGGAGAGCTGAAATATGACTGTGTGAACCTGATT	1020		
DB	1037	GAGAAATTCACGACTATCTCTATTCGGAGAGCTGAAATATGACTGTGTGTGAACCTGATT	1096		
QY	1021	TCCACCGGAGGTGACAGAGCCCACTTACGAGCTTGTGAATCACCAAAATGCTG	1080		

Db 1097 TCCACCGGAGTGTGAGGCCCACTACTACGGCTGTGAAGTCACCAACCAATGCTG 1156
Qy 1081 TGTGCTGTGACCCACAGTGTGAAAAAGATTCCTGSCCAGAGAGACTAGGGGAGCCCTC 1140
Db 1157 TGTGCTGTGACCCACAGTGTGAAAAAGATTCCTGSCCAGAGAGACTAGGGGAGCCCTC 1216
Qy 1141 GTCTGTTCCTTCCAAAGGCGCGCATGACTTTGACTGGAATTGTAGCTGGGGCCGTGATGT 1200
Db 1217 GTCTGTTCCTTCCAAAGGCGCGCATGACTTTGACTGGAATTGTAGCTGGGGCCGTGATGT 1276
Qy 1201 GCCCTGAAGACAAGCCAGGCGTCTACACAGAGAGTCTACACTTACCTTGAATCCGC 1260
Db 1277 GCCCTGAAGACAAGCCAGGCGTCTACACAGAGAGTCTACACTTACCTTGAATCCGC 1336
Qy 1261 AGTCACACCAAGAAAGAGATGGCTGGCCCTCTGA 1296
Db 1337 AGTCACACCAAGAAAGAGATGGCTGGCTCTCTGA 1372

RESULT 14

US-07-957-039A-7
; Sequence 7, Application US/07957039A
; Patent No. 5389538
; GENERAL INFORMATION:
; APPLICANT: TANABE, TOSHIKIZUMI
; APPLICANT: MORITA, MASANORI
; APPLICANT: HIROSE, MASAKI
; APPLICANT: AMASUJI, YASUO
; TITLE OF INVENTION: MUTANT HUMAN PROUKINASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/957,039A
; FILING DATE: 06-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 289257/1991
; FILING DATE: 07-OCT-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1233
US-07-957-039A-7

Query Match 95.1%; Score 1232.8; DB 1; Length 1236;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 61 AGCAATGAAGTTCATCAAGTTCATGCAAGTGTGACTGTCTAATGAGAGAAACATGTGTG 120

Db 1 AGCAATGAAGTTCATCAAGTTCATGCAAGTGTGACTGTCTAATGAGAGAAACATGTGTG 60
Qy 121 TCCAAAGTACTTCTCCCAACATTCAGTGTGCAACTGCCCCAAGAAATTTGGAGGGCAG 180
Db 61 TCCAAAGTACTTCTCCCAACATTCAGTGTGCAACTGCCCCAAGAAATTTGGAGGGCAG 120
Qy 181 CACTGTGAATAGATTAAGTCAAAAACCTGCTATGAGGGAAATGTCACTTTTACCGAGGA 240
Db 121 CACTGTGAATAGATTAAGTCAAAAACCTGCTATGAGGGAAATGTCACTTTTACCGAGGA 180
Qy 241 AAGGCTCAGACTGACACCACTGGGCGGCCCTGTGCTCTGCAACTTGGCACTGTCTT 300
Db 181 AAGGCTCAGACTGACACCACTGGGCGGCCCTGTGCTCTGCAACTTGGCACTGTCTT 240
Qy 301 CAGCAAAAGTACATCATGCCCAAGATCTGATGCTCTTGAAGCTGGGGCAACATAT 360
Db 241 CAGCAAAAGTACATCATGCCCAAGATCTGATGCTCTTGAAGCTGGGGCAACATAT 300
Qy 361 TACTCAGGAACCCAGACAACCGAGGCGACCTGTGTCTATGTGCAAGTGGGCTTAAAG 420
Db 301 TACTCAGGAACCCAGACAACCGAGGCGACCTGTGTCTATGTGCAAGTGGGCTTAAAG 360
Qy 421 CGGCTTGTCCAAAGATGTGATGTGATGATGATGATGATGATGATGATGATGATGATG 480
Db 361 CGGCTTGTCCAAAGATGTGATGTGATGATGATGATGATGATGATGATGATGATGATG 420
Qy 481 CCAGAAAGATTTAAATTTCAAGTGTGCAAAAAGACTGAGGGCCCGCTTTAAGATAT 540
Db 421 CCAGAAAGATTTAAATTTCAAGTGTGCAAAAAGACTGAGGGCCCGCTTTAAGATAT 480
Qy 541 GGGGGAAGATTCACCAACCATGTGAACCAAGCTTGTGTGAGGCACTTACAGAGGAC 600
Db 481 GGGGGAAGATTCACCAACCATGTGAACCAAGCTTGTGTGAGGCACTTACAGAGGAC 540
Qy 601 CGGGGGGGCTGTGACCTAGTGTGTGAGGAGGAGCCCTCAATGAGCCCTTGTGGTATC 660
Db 541 CGGGGGGGCTGTGACCTAGTGTGTGAGGAGGAGCCCTCAATGAGCCCTTGTGGTATC 600
Qy 661 AGCGCACACACTGCTTATGATTAACCAAGAAAGAGAGCTACATGCTTACCTGGGT 720
Db 601 AGCGCACACACTGCTTATGATTAACCAAGAAAGAGAGCTACATGCTTACCTGGGT 660
Qy 721 CGCTCAAGGCTTAATCTCAACACCAAGGGAATGAAGTTGAGTGTGAAAACTTATC 780
Db 661 CGCTCAAGGCTTAATCTCAACACCAAGGGAATGAAGTTGAGTGTGAAAACTTATC 720
Qy 781 CTACACAAGGACTACAGGCTGACAGGCTTGTCTACACCAACGACATTTGCTGTGAAG 840
Db 721 CTACACAAGGACTACAGGCTGACAGGCTTGTCTACACCAACGACATTTGCTGTGAAG 780
Qy 841 ATCGGTTCCAAAGAGGAGGTGTGCGACAGCCATCCGGAATATACAGACCATCTGCTG 900
Db 781 ATCGGTTCCAAAGAGGAGGTGTGCGACAGCCATCCGGAATATACAGACCATCTGCTG 840
Qy 901 CCTCGATGATTAAGATTCCTCCAGTTTGGCAAGCTGTGATGATGATGATGATGATGATG 960
Db 841 CCTCGATGATTAAGATTCCTCCAGTTTGGCAAGCTGTGATGATGATGATGATGATGATG 900
Qy 961 GAGAATTTCAACGACTATCTTATCCGAGACAGTGAATAATGACTGTGGAAGTGAAT 1020
Db 901 GAGAATTTCAACGACTATCTTATCCGAGACAGTGAATAATGACTGTGGAAGTGAAT 960
Qy 1021 TCCCAACGGGAGTGTGACAGGCCCACTACTACAGGCTGTGAAGTGTACCAACAAATGCTG 1080
Db 961 TCCCAACGGGAGTGTGACAGGCCCACTACTACAGGCTGTGAAGTGTACCAACAAATGCTG 1020
Qy 1081 TGTGCTGTGACCCACAGTGTGAAAAAGATTCCTGSCCAGAGAGACTAGGGGAGCCCTC 1140
Db 1021 TGTGCTGTGACCCACAGTGTGAAAAAGATTCCTGSCCAGAGAGACTAGGGGAGCCCTC 1080
Qy 1141 GTCTGTTCCTTCCAAAGGCGCGCATGACTTTGACTGGAATTGTGAGCTGGGGCCGTGATGT 1200

Db 1081 GTCGTCTCCCTCCAGGCGCATGACTTGTGATGAAATTGTGAGCTGGGCGGTGATGT 1140
QY 1201 GCCCTAAGAGCAAGCAGGCGCTCTACAGAGACTCTCACTTTTACCTTGGAATCCGC 1260
Db 1141 GCCCTAAGAGCAAGCAGGCGCTCTACAGAGACTCTCACTTTTACCTTGGAATCCGC 1200
QY 1261 AGTCACCAAGAGAGAGATGAGCTGGCCCTCTGA 1296
Db 1201 AGTCACCAAGAGAGAGATGAGCTGGCCCTCTGA 1236

RESULT 15
US-09-023-655-927
; Sequence 927, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 927:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 91311467
; US-09-023-655-927

Query Match 95.1%; Score 1232.8; DB 4; Length 1236;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 61 AGCAATGAATCTCATCAAGTTCATGCACTGTAAGTCTTAATGAGAAATGTGTG 120
Db 1 AGCAATGAATCTCATCAAGTTCATGCACTGTAAGTCTTAATGAGAAATGTGTG 60
QY 121 TCACAACAGTACTTCTCCCAACATTCACCTGTGCACTGCCCAAGAAATTCGAGGCGAG 180
Db 61 TCACAACAGTACTTCTCCCAACATTCACCTGTGCACTGCCCAAGAAATTCGAGGCGAG 120
QY 181 CACTGTGAATAATGATTAAGTCAAAACCTGTCTATGAGGGAGATGTCACCTTTTACGAGAGA 240

Db 121 CACTGTGAATAATGATTAAGTCAAAACCTGTCTATGAGGGAGATGTCACCTTTTACGAGAGA 180
QY 241 AAGGCGAGACATGACACCAATGAGCGCGCTGCTGCTCCCTGGAATCTGTGCACTGTCTT 300
Db 181 AAGGCGAGACATGACACCAATGAGCGCGCTGCTGCTCCCTGGAATCTGTGCACTGTCTT 240
QY 301 CAGCAACGATACCATGCGCCACAGATGATGCTCTTCACTGGGCGCTGGGGAACATTAAT 360
Db 241 CAGCAACGATACCATGCGCCACAGATGATGCTCTTCACTGGGCGCTGGGGAACATTAAT 300
QY 361 TACTGAGAGAACCCAGACCAACCGAGCGCACCTGTGCTATGTGAGGTGAGGCTTAAG 420
Db 301 TACTGAGAGAACCCAGACCAACCGAGCGCACCTGTGCTATGTGAGGTGAGGCTTAAG 360
QY 421 CGGCTTGTCCAAAGATGATGCTGATGATGCTGCGCATGAGAAAGCCCTCTCTCTCT 480
Db 361 CTGCTTGTCCAAAGATGATGCTGATGATGCTGCGCATGAGAAAGCCCTCTCTCTCT 420
QY 481 CCAGAAATTTAAATTTGATGCTGCGCAAAAGACTCTGAGCGCCCTTTAAGATTAAT 540
Db 421 CCAGAAATTTAAATTTGATGCTGCGCAAAAGACTCTGAGCGCCCTTTAAGATTAAT 480
QY 541 GGGGAGAAATTCACCAACCATGAGAACAGCCCTGTGCTTGGCGCATCTACAGAGAGCAC 600
Db 481 GGGGAGAAATTCACCAACCATGAGAACAGCCCTGTGCTTGGCGCATCTACAGAGAGCAC 540
QY 601 CGGGGGGGCTGTCTACCTGATGCTGAGGAGCGAGCTTCATACGCCCTTGTGGGTATC 660
Db 541 CGGGGGGGCTGTCTACCTGATGCTGAGGAGCGAGCTTCATACGCCCTTGTGGGTATC 600
QY 661 AGCGCCACACTGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
Db 601 AGCGCCACACTGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
QY 721 CGCTCAAGGCTTAATCTCAACAGCAAGGAGAGATGAGTTGAGGTGAGAAACCTAATC 780
Db 661 CGCTCAAGGCTTAATCTCAACAGCAAGGAGAGATGAGTTGAGGTGAGAAACCTAATC 720
QY 781 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACACAGCAAGCAATTCCTTGTGAAG 840
Db 721 CTACACAAAGACTACAGCGCTGACAGCTTGTCTACACAGCAAGCAATTCCTTGTGAAG 780
QY 841 ATCCGTTCCAAAGAGAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 781 ATCCGTTCCAAAGAGAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 901 CCTCGATGTATTAAGATCCCGAGTTTGTGACAAAGCTGTGAGATCACTGGCTTTGAGAAA 960
Db 841 CCTCGATGTATTAAGATCCCGAGTTTGTGACAAAGCTGTGAGATCACTGGCTTTGAGAAA 900
QY 961 GAGAAATTTACCGACTATCTATCCGAGAGAGCTGAGAAATGACTGTGTGAGAGCTGATT 1020
Db 901 GAGAAATTTACCGACTATCTATCCGAGAGAGCTGAGAAATGACTGTGTGAGAGCTGATT 960
QY 1021 TCCACACCGGAGATGTGAGAGAGCCCACTACTACGCTCTGAAAGTCAACCAAAATGCTG 1080
Db 961 TCCACACCGGAGATGTGAGAGAGCCCACTACTACGCTCTGAAAGTCAACCAAAATGCTG 1020
QY 1081 TGTGCTGTGACCCACAGTGAAGAAACAGATTTCTGTGCAAGGAGACTCAAGGAGAGACCCCTC 1140
Db 1021 TGTGCTGTGACCCACAGTGAAGAAACAGATTTCTGTGCAAGGAGACTCAAGGAGAGACCCCTC 1080
QY 1141 GTCTGTTCCTTCCAAAGCGCGCATGACTTGAATGAGTGTGAGTGTGAGGCGCTGTGATGT 1200
Db 1081 GTCTGTTCCTTCCAAAGCGCGCATGACTTGAATGAGTGTGAGTGTGAGGCGCTGTGATGT 1140
QY 1201 GCGCTGAAGAGACAAAGCGAGGCTTACACAGAGATCTCACTTTTACCTTGATCCGC 1260
Db 1141 GCGCTGAAGAGACAAAGCGAGGCTTACACAGAGATCTCACTTTTACCTTGATCCGC 1200
QY 1261 AGTCACCAAGAGAGAGATGAGCTGGCCCTCTGA 1296

Db 1201 AGTCACCAAGAGAGATGGCTGGCCCTGA 1236

Search completed: March 21, 2004, 20:52:09
Job time : 128 secs


```

Oy 481 CCAGAAAGATTAAATTTAGTGTGGCCAAAAGACTCTGAGGCCCGCTTTAAGATTATT 540
Db |||||
Oy 547 CCAGAAAGATTAAATTTAGTGTGGCCAAAAGACTCTGAGGCCCGCTTTAAGATTATT 606
Db |||||
Oy 541 GGGGGAGAAATTACCAACCACTGAGAAACCAAGCCCTGTGTTGGCGGCATCTACAGAGGCAC 600
Db GGGGGAGAAATTACCAACCACTGAGAAACCAAGCCCTGTGTTGGCGGCATCTACAGAGGCAC 666
Oy 601 GGGGGGGGCTCTGTCACTTACGTTGTGAGAGGCAAGCTTCATCAGCCCTTGTGGTGTATC 660
Db GGGGGGGGCTCTGTCACTTACGTTGTGAGAGGCAAGCTTCATCAGCCCTTGTGGTGTATC 726
Oy 661 AGCGGCACACACTGCTTCAATTGATTTACCCAAAGAAAGAGAGACTACATTCGTCTACCTGGGT 720
Db AGCGGCACACACTGCTTCAATTGATTTACCCAAAGAAAGAGAGACTACATTCGTCTACCTGGGT 786
Oy 721 CGCTCAAGGCTTAACTCCAAACGCAAGAGGGAGATGAACTTTGAGGTGGAAGAAACCTAAATC 780
Db CGCTCAAGGCTTAACTCCAAACGCAAGAGGGAGATGAACTTTGAGGTGGAAGAAACCTAAATC 846
Oy 781 CTACCAAGAGACTACAGCGCTGACACGCTTGTCTACCAACAACGACATTTGCTTGTGAAG 840
Db CTACCAAGAGACTACAGCGCTGACACGCTTGTCTACCAACAACGACATTTGCTTGTGAAG 906
Oy 847 CTACCAAGAGACTACAGCGCTGACACGCTTGTCTACCAACAACGACATTTGCTTGTGAAG 906
Db CTACCAAGAGACTACAGCGCTGACACGCTTGTCTACCAACAACGACATTTGCTTGTGAAG 966
Oy 841 ATCCGTTCCAAAGAGGGGAGGTGTGGGAGCCATCCCGGACTTAAAGACATCTGCTG 900
Db ATCCGTTCCAAAGAGGGGAGGTGTGGGAGCCATCCCGGACTTAAAGACATCTGCTG 966
Oy 901 CCCTCGATGTATATAGCATCCCAAGTTTGGCACAGCTGTGAGATCACTGGCTTTGGAAAA 960
Db CCCTCGATGTATATAGCATCCCAAGTTTGGCACAGCTGTGAGATCACTGGCTTTGGAAAA 102
Oy 961 GAGAAATTCTACCGACTATTTCTATCCGAGCAGCTGAAATGACTGTGTGAAGCTGATT 102
Db GAGAAATTCTACCGACTATTTCTATCCGAGCAGCTGAAATGACTGTGTGAAGCTGATT 108
Oy 1021 TCCCAACGGGAGTGTACAGAGGCCCACTATCCGCTCGAAGTCAACCAAAATGCTG 108
Db TCCCAACGGGAGTGTACAGAGGCCCACTATCCGCTCGAAGTCAACCAAAATGCTG 114
Oy 1081 TGTGTCTGTGACCCACAGTGTGAAAAAGATTCTCTGCACAGGAGACTCAGGGGACCCCTC 114
Db TGTGTCTGTGACCCACAGTGTGAAAAAGATTCTCTGCACAGGAGACTCAGGGGACCCCTC 120
Oy 1141 GTCTGTTCCTCTCCAAAGCCCGCATGACTTTGACTGGAATTGTGAGCTGTGGAGCGTGTATGT 120
Db GTCTGTTCCTCTCCAAAGCCCGCATGACTTTGACTGGAATTGTGAGCTGTGGAGCGTGTATGT 126
Oy 1201 GCCCGGAAGGACAAACGAGGCGCTTACACAGAGAGTCAACATTCCTTACCCGTGATCGCG 126
Db GCCCGGAAGGACAAACGAGGCGCTTACACAGAGAGTCAACATTCCTTACCCGTGATCGCG 132
Oy 1261 AGTCACACCAAGAAAGAGATGGCTGTGGCCCTCTGA 1296
Db AGTCACACCAAGAAAGAGATGGCTGTGGCCCTCTGA 1362

RESULT 3
US-09-850-716A-123
; Sequence 123, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0

```

Query Match	100.0%	Score 1296;	DB 9;	Length 2294;
Beet Local Similarity	100.0%	Pred. No. 0;		
Matches 1296;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
SEQ ID NO 123				
LENGTH: 2294				
TYPE: DNA				
ORGANISM: Homo sapien				
US-09-850-716A-123				
QY	1	ATGAGAGCCCTGCTGGCGCGCCCTGCTTCTCTGCGCTGCTGTCGTGTAGCGACTCCAAAGC	60	
Db	67	ATGAGAGCCCTGCTGGCGCGCCCTGCTTCTCTGCGCTGCTGTCGTGTAGCGACTCCAAAGC	126	
QY	61	AGCAATGAATCTTCATCAAGTTCATCGAATCTGTGACTGTCTTAATGAGAGAAATGTGTG	120	
Db	127	AGCAATGAATCTTCATCAAGTTCATCGAATCTGTGACTGTCTTAATGAGAGAAATGTGTG	186	
QY	121	TTCAACAAAGTACTTCTCCAACTTCACCTGTGTGCAATCGCCCAAGAAATTTGGAGGGCAG	180	
Db	187	TTCAACAAAGTACTTCTCCAACTTCACCTGTGTGCAATCGCCCAAGAAATTTGGAGGGCAG	246	
QY	181	CACGTGAAATAGATTAAGTCAAAAAACCGTATGAGGGGAATGGCACTTTTAAACGAGAGA	240	
Db	247	CACGTGAAATAGATTAAGTCAAAAAACCGTATGAGGGGAATGGCACTTTTAAACGAGAGA	306	
QY	241	AAGGCAAGCACTGACACCCATGGGCGGCGCTGCTGCTGCTGGAATCTTGCACTGTCTT	300	
Db	307	AAGGCAAGCACTGACACCCATGGGCGGCGCTGCTGCTGCTGGAATCTTGCACTGTCTT	366	
QY	301	CAGCAAAAGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	360	
Db	367	CAGCAAAAGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	426	
QY	361	TACTGCAAGAAACCCAGACAAACCGGAGGCGACCCGTGATGATGATGATGATGATGATGATG	420	
Db	427	TACTGCAAGAAACCCAGACAAACCGGAGGCGACCCGTGATGATGATGATGATGATGATGATG	486	
QY	421	CCGCTTGTCCAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	480	
Db	487	CCGCTTGTCCAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	546	
QY	481	CCAGAAAGATTTAAATTTCACTGTGCGCAAAAGACTCTGAGCCCGCTTTAAATTTAT	540	
Db	547	CCAGAAAGATTTAAATTTCACTGTGCGCAAAAGACTCTGAGCCCGCTTTAAATTTAT	606	
QY	541	GGGGGGAATTCACCAACCATGGAACCAAGCCCTGTGTTGGGGCATTTACAGAGAGAC	600	
Db	607	GGGGGGAATTCACCAACCATGGAACCAAGCCCTGTGTTGGGGCATTTACAGAGAGAGAC	666	
QY	601	CGGGGGGGCTGTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	660	
Db	667	CGGGGGGGCTGTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	726	
QY	661	AGCGCAACACACTGCTTCAATTTATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720	
Db	727	AGCGCAACACACTGCTTCAATTTATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	786	
QY	721	CGCTCAAGGCTTAACTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	780	
Db	787	CGCTCAAGGCTTAACTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	846	
QY	781	CTACACAAAGAGCTACAGGCTGACAGCGTTGCTACCAACAGAGAGAGAGAGAGAGAGAGAG	840	
Db	847	CTACACAAAGAGCTACAGGCTGACAGCGTTGCTACCAACAGAGAGAGAGAGAGAGAGAGAG	906	
QY	841	ATCGGTTTCAAG	900	
Db	907	ATCGGTTTCAAG	966	
QY	901	CCCTTCATGATTAACGATCCCAAGTTTGGCAAGAGCTGTGATGATGATGATGATGATGATGAT	960	
Db	967	CCCTTCATGATTAACGATCCCAAGTTTGGCAAGAGCTGTGATGATGATGATGATGATGATGAT	1026	

```
QY 961 GAGATTCTACCGACTATCTATCCGAGACGTGAAAATGACTTTGTGAACTGATT 1020
Db 1027 GAGAAATCTACCGACTATCTATCCGAGAGCGTGAATAATGACTGTTGTGAAGTATT 1086
QY 1021 TCCACCGGAGTGTCAAGAGCCCACTACTACGGCTGTGAAGTACCAACCAAAATGCTG 1080
Db 1087 TCCACCGGAGTGTCAAGAGCCCACTACTACGGCTGTGAAGTACCAACCAAAATGCTG 1146
QY 1081 TGTGCTGTGACCCACAGTGTGAAAACAGATTCTGTCCAGGGAGACTAGGGGGAGCCCTC 1140
Db 1147 TGTGCTGTGACCCACAGTGTGAAAACAGATTCTGTCCAGGGAGACTAGGGGGAGCCCTC 1206
QY 1141 GTCTGTTCCCTCCAGAGCCGATGACTTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1200
Db 1207 GTCTGTTCCCTCCAGAGCCGATGACTTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1266
QY 1201 GCCCTGAAGACAAAGCCAGGCGTCTACAGAGTCTACACTTTACCTGATCCG 1260
Db 1267 GCCCTGAAGACAAAGCCAGGCGTCTACAGAGTCTACACTTTACCTGATCCG 1326
QY 1261 AGTACACCAAGAGAAAGAAATGGCTGGCCCTCTGA 1296
Db 1327 AGTACACCAAGAGAAAGAAATGGCTGGCCCTCTGA 1362
```

```
RESULT 4
US-09-897-778-123
; Sequence 123. Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedivick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Matanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-123
```

```
Query Match 100.0%; Score 1296; DB 9; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGAGAGCCCTGTGCGCGGCTGTCTCTGTGCTGCTGTGAGTGTGAGTGTGAGTGTGAGTGT 60
Db 67 ATGAGAGCCCTGTGCGCGGCTGTCTCTGTGCTGCTGTGAGTGTGAGTGTGAGTGTGAGTGT 126
QY 61 AGCAATGAAGCTTCATCAATGATTCATGAACTGTGACTGTCTAAATGAGGAAATCATGTGTG 120
Db 127 AGCAATGAAGCTTCATCAATGATTCATGAACTGTGACTGTCTAAATGAGGAAATCATGTGTG 186
QY 121 TCCAAAGTACTTCTTCCAAATTCATGCTGTGCAACTGCCAAAGAAATTTGGAGGGGAG 180
Db 187 TCCAAAGTACTTCTTCCAAATTCATGCTGTGCAACTGCCAAAGAAATTTGGAGGGGAG 246
QY 181 CACTGTGAATAGATAAGTCAAAAACCTGTCTATGAGGGGAATGATCACTTTTACCGAGA 240
Db 247 CACTGTGAATAGATAAGTCAAAAACCTGTCTATGAGGGGAATGATCACTTTTACCGAGA 306
```

```
QY 241 AAGGCCAGCACTGACACCATGAGCGGCGGCTGCTGCGCTGGAACTGTGCCATGTCTCT 300
Db 307 AAGGCCAGCACTGACACCATGAGCGGCGGCTGCTGCGCTGGAACTGTGCCATGTCTCTCT 366
QY 301 CAGCAAACTTACATGATCCCAAGATCTGATGTCTTCAAGTGTGGGCTTGGGAAACATAT 360
Db 367 CAGCAAACTTACATGATCCCAAGATCTGATGTCTTCAAGTGTGGGCTTGGGAAACATAT 426
QY 361 TACTGAGGAACCCAGACAAACCGAGGAGGAGCCCTGTGTGTATGTGACAGTGTGGGCTTAAG 420
Db 427 TACTGAGGAACCCAGACAAACCGAGGAGGAGCCCTGTGTGTATGTGACAGTGTGGGCTTAAG 486
QY 421 CCGCTTTGTCAAAGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 487 CCGCTTTGTCAAAGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
QY 481 CCAGAAAGATTTAAATTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 547 CCAGAAAGATTTAAATTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 606
QY 541 GGGGAGAAATTCACCAACATGAGAACAGAGCCCTGTTTGGGCGATCTACAGAGGAGC 600
Db 607 GGGGAGAAATTCACCAACATGAGAACAGAGCCCTGTTTGGGCGATCTACAGAGGAGC 666
QY 601 GGGGGGGGCTGTGTCACTTACGTGTGAGAGGAGAGCTCATGAGCCCTTGTGTGTGATC 660
Db 667 GGGGGGGGCTGTGTCACTTACGTGTGAGAGGAGAGCTCATGAGCCCTTGTGTGTGATC 726
QY 661 AGCGCCACACTGCTTATGATTTACCAAGAAAGAGAGATCAATGTCTACCTGAGT 720
Db 727 AGCGCCACACTGCTTATGATTTACCAAGAAAGAGAGATCAATGTCTACCTGAGT 786
QY 721 GCTCTAAGCTTAACTTCAACAGCAAGGAGAGATGAAGTTTGAAGTGAAGAACTTAATC 780
Db 787 GCTCTAAGCTTAACTTCAACAGCAAGGAGAGATGAAGTTTGAAGTGAAGAACTTAATC 846
QY 781 CTACACAAGGACTACAGGCTGTGACAGCTGTCTACCAACCAACATGTGCTGTGAAG 840
Db 847 CTACACAAGGACTACAGGCTGTGACAGCTGTCTACCAACCAACATGTGCTGTGAAG 906
QY 841 ATTCGTTCCAAAGAGGAGGAGGTGTGCGAGCAATCCCGAGCTATACAGACATCTGCTG 900
Db 907 ATTCGTTCCAAAGAGGAGGAGGTGTGCGAGCAATCCCGAGCTATACAGACATCTGCTG 966
QY 901 CCTCCGATGTATTAACGATCCCGAGTTTGGCAAGCTGTGATCACTGCTTTGAGAAA 960
Db 967 CCTCCGATGTATTAACGATCCCGAGTTTGGCAAGCTGTGATCACTGCTTTGAGAAA 1026
QY 961 GAGAAATCTACCGACTATCTATCCGAGACGTGAAAATGACTGTGTGAAAGTGAAT 1020
Db 1027 GAGAAATCTACCGACTATCTATCCGAGACGTGAAAATGACTGTGTGAAAGTGAAT 1086
QY 1021 TCCACCGGAGTGTCAAGAGCCCACTACTACGGCTGTGAAGTACCAACCAAAATGCTG 1080
Db 1087 TCCACCGGAGTGTCAAGAGCCCACTACTACGGCTGTGAAGTACCAACCAAAATGCTG 1146
QY 1081 TGTGCTGTGACCCACAGTGTGAAAACAGATTCTGTCCAGGGAGACTAGGGGGAGCCCTC 1140
Db 1147 TGTGCTGTGACCCACAGTGTGAAAACAGATTCTGTCCAGGGAGACTAGGGGGAGCCCTC 1206
QY 1141 GTCTGTTCCCTCCAGAGCCGATGACTTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1200
Db 1207 GTCTGTTCCCTCCAGAGCCGATGACTTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1266
QY 1201 GCCCTGAAGACAAAGCCAGGCGTCTACAGAGTCTACACTTTTACCTGATCCG 1260
Db 1267 GCCCTGAAGACAAAGCCAGGCGTCTACAGAGTCTACACTTTTACCTGATCCG 1326
QY 1261 AGTACACCAAGAGAAAGAAATGGCTGGCCCTCTGA 1296
Db 1327 AGTACACCAAGAGAAAGAAATGGCTGGCCCTCTGA 1362
```

```
RESULT 5
US-09-466-396A-123
; Sequence 123, Application US/09466396A
; Publication No. US20030119763A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-123

Query Match      100.0%; Score 1296; DB 10; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGCGCGCCCTGCTTCTCTGCTCCTGCTGAGCACTCCAAAGC 60
DB 67 ATGAGAGCCCTGCTGCGCGCCCTGCTTCTCTGCTCCTGCTGAGCACTCCAAAGC 126
QY 61 AGCAATGAATCTTCATCAAGTTCCTCATGAACTGTGACTGTCTAAATGAGAGAACTGTGTG 120
DB 127 ACCAATGAATCTTCATCAAGTTCCTCATGAACTGTGACTGTCTAAATGAGAGAACTGTGTG 186
QY 121 TCCAAACAATCTTCTCCCAACATTCACGTGTGCAATGCGCCAAAGAAATTCGAGGGCAG 180
DB 187 TCCAAACAATCTTCTCCCAACATTCACGTGTGCAATGCGCCAAAGAAATTCGAGGGCAG 246
QY 181 CACTGTGAATAGATAGTCAAAACCTGCTATGAGGGAGATGTCATCTTTTACCGAGA 240
DB 247 CACTGTGAATAGATAGTCAAAACCTGCTATGAGGGAGATGTCATCTTTTACCGAGA 306
QY 241 AAGGCCAGACTGACACCATGAGGCGCCCTGCTGCTGCTGCAACTCTGCTCTCT 300
DB 307 AAGGCCAGACTGACACCATGAGGCGCCCTGCTGCTGCTGCAACTCTGCTCTCTCT 366
QY 301 CAGCAAAAGTACATGAGCCCAAGATGCTGCTCTTCACTGAGGCTGAGGAAACATAT 360
DB 367 CAGCAAAAGTACATGAGCCCAAGATGCTGCTCTTCACTGAGGCTGAGGAAACATAT 426
QY 361 TACTCAGGAACCCAGACAAACCGAGAGCCACCTGCTGCTATGTCAGAGGCTCTAAAG 420
DB 427 TACTCAGGAACCCAGACAAACCGAGAGCCACCTGCTGCTATGTCAGAGGCTCTAAAG 486
QY 421 CCGCTTGTTCAGAGAGTGCATGCTGTCAGTCTGCGCAGATGAGAAAAAGCCCTCTCTCT 480
DB 487 CCGCTTGTTCAGAGAGTGCATGCTGTCAGTCTGCGCAGATGAGAAAAAGCCCTCTCTCT 546
QY 481 CCAGAAAGATTAATTTTCAAGTGTGCGCCAAAGAGCTCTAGAGCCCGCTTTTAAGATTAT 540
DB 547 CCAGAAAGATTAATTTTCAAGTGTGCGCCAAAGAGCTCTAGAGCCCGCTTTTAAGATTAT 606
QY 541 GGGGGAGAAATTCACCAACCATGAGAACCGCCCTGCTGCTGCGCCATCTACAGAGAGCAC 600
DB 607 GGGGGAGAAATTCACCAACCATGAGAACCGCCCTGCTGCTGCGCCATCTACAGAGAGCAC 666
QY 601 CGGGGGGGCTGTGCACTGCTGTGTGTGAGGCAAGCTTCATCAGCCCTTGTCTGCTGATC 660
DB 667 CGGGGGGGCTGTGCACTGCTGTGTGTGAGGCAAGCTTCATCAGCCCTTGTCTGCTGATC 726
QY 661 AGCGCCACACATGCTCTTCAATTAATTAACCAAGAGAGAGATTAATCTCTACCTGGGT 720
DB 727 AGCGCCACACATGCTCTTCAATTAATTAACCAAGAGAGAGATTAATCTCTACCTGGGT 786
QY 721 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAACTTTGAGGTGAGAAACCTAATC 780
```

```
DB 787 CGCTCAAGGCTTAATCTCAACACGCAAGGGAGATGAACTTTGAGGTGAGAAACCTAATC 846
QY 781 CTACACAAAGACTACAGCGCGCTGACAGGCTTGTCTACACACACACATTTGCTTGTGAAG 840
DB 847 CTACACAAAGACTACAGCGCGCTGACAGGCTTGTCTACACACACACATTTGCTTGTGAAG 906
QY 841 ATCCGTTCCAAAGAGGGAGGTGTGCGAGCCATCCCGGACTTAACAGACATCTGCTG 900
DB 907 ATCCGTTCCAAAGAGGGAGGTGTGCGAGCCATCCCGGACTTAACAGACATCTGCTG 966
QY 901 CCCTGATGTATTAACGATCCCAAGTTTGGCACAAAGCTGTGAGATCACTGCTTTGGAAAA 960
DB 967 CCTGATGTATTAAGATCCCAAGTTTGGCACAAAGCTGTGAGATCACTGCTTTGGAAAA 1026
QY 961 GAGAAATTCACGACTATCTTATCCGAGCAGCTGAAATGACTGTTGTGAAGTGAAT 1020
DB 1027 GAGAAATTCACGACTATCTTATCCGAGCAGCTGAAATGACTGTTGTGAAGTGAAT 1086
QY 1021 TCCCAACGGGAGGTGAGAGAGCCCACTACTACGGCTGTAAGTCAACCAAAATGCTG 1080
DB 1087 TCCCAACGGGAGGTGAGAGAGCCCACTACTACGGCTGTAAGTCAACCAAAATGCTG 1146
QY 1081 TGTGCTGTGACCCACAGTGAAGAAACAGATTCCTGCGCAGGAGACTCAGGGGAGCCCTC 1140
DB 1147 TGTGCTGTGACCCACAGTGAAGAAACAGATTCCTGCGCAGGAGACTCAGGGGAGCCCTC 1206
QY 1141 GTTGTGCTTCCCTCCAAAGCCGCAATGACTTGTGACTGTAATGTGAGCTGGGGCCGTGATGT 1200
DB 1207 GTTGTGCTTCCCTCCAAAGCCGCAATGACTTGTGACTGTAATGTGAGCTGGGGCCGTGATGT 1266
QY 1201 GCCCTGAAGAGCAAGCCGCTTACAGAGAGTGCACACTTCACTCCGTGAATCCG 1260
DB 1267 GCCCTGAAGAGCAAGCCGCTTACAGAGAGTGCACACTTCACTCCGTGAATCCG 1326
QY 1261 AGTCACACCCAAAGAGAGATGCGCTGCGCTCTGA 1296
DB 1327 AGTCACACCCAAAGAGAGATGCGCTGCGCTCTGA 1362

RESULT 6
US-10-117-982-123
; Sequence 123, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Matanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; NUMBER OF SEQ ID NOS: 2002-04-05
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-982-123

Query Match      100.0%; Score 1296; DB 14; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGCTGCGCGCCCTGCTTCTCTGCTCCTGCTGAGCACTCCAAAGC 60
```

Db 67 ATGAGAGCCCTGCTGGGCGGCTCTCTCTGCGCTCTGCTGCGAGCGACTCCAAAGGC 126
Qy 61 AGCAATGAACCTTCACTCAAGTTCATCGCAATGTGACTGTCTAAATVGAGAGAACATGTGTG 120
Db 127 AGCAATGAACCTTCACTCAAGTTCATCGCAATGTGACTGTCTAAATVGAGAGAACATGTGTG 186
Qy 121 TCCAAACAAGTACTTCTCCAAATTCATCTGTGTGCACTGTGCCCAAGAAATTCGAGGGCAG 180
Db 187 TCCAAACAAGTACTTCTCCAAATTCATCTGTGTGCACTGTGCCCAAGAAATTCGAGGGCAG 246
Qy 181 CACTGTGAATATAGTAAAGTCAAAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGGA 240
Db 247 CACTGTGAATATAGTAAAGTCAAAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGGA 306
Qy 241 AAGGCCAGCACTGACACCAATGGGCGGCGCTGCTGCTGGAACCTGTGCACTGTCTCTT 300
Db 307 AAGGCCAGCACTGACACCAATGGGCGGCGCTGCTGCTGGAACCTGTGCACTGTCTCTT 366
Qy 301 CAGCAACGTAACCACTGCGCCACAGATCTGATGCTCTTACGTGGGCGCTGGGGAACATTAAT 360
Db 367 CAGCAACGTAACCACTGCGCCACAGATCTGATGCTCTTACGTGGGCGCTGGGGAACATTAAT 426
Qy 361 TACTGCAGGAACCCAGCAACCCGAGGCGAACCTGTGTCTATGTGCAAGTGGGCTTAAAG 420
Db 427 TACTGCAGGAACCCAGCAACCCGAGGCGAACCTGTGTCTATGTGCAAGTGGGCTTAAAG 486
Qy 421 CCGCTGTCCAAAGTGCATGTGTGATGATCTGGCGCAATGGAAAAAGCCCTCTCTCTCT 480
Db 487 CCGCTGTCCAAAGTGCATGTGTGATGATCTGGCGCAATGGAAAAAGCCCTCTCTCTCT 546
Qy 481 CCAGAAAGATTTAAATTTCAAGTGTGGGCAAAAGCTCTGAGGCGCGCTTTAAGATTATT 540
Db 547 CCAGAAAGATTTAAATTTCAAGTGTGGGCAAAAGCTCTGAGGCGCGCTTTAAGATTATT 606
Qy 541 GGGGAGAAATTCACACCATCGAGAACCAAGCCCTGTGTTGGGCGCATCTTACAGAGGACAC 600
Db 607 GGGGAGAAATTCACACCATCGAGAACCAAGCCCTGTGTTGGGCGCATCTTACAGAGGACAC 666
Qy 601 CGGGGGGGCTCTGCTCACTGCTGTGTGAGGAGCGCTCATACGCCCTTGTGGGCTGATG 660
Db 667 CGGGGGGGCTCTGCTCACTGCTGTGTGAGGAGCGCTCATACGCCCTTGTGGGCTGATG 726
Qy 661 AGCGCCACACACTGCTTCAATTTATTAACCAAGAGAGAGACTTACATCTGCTGCTGCTG 720
Db 727 AGCGCCACACACTGCTTCAATTTATTAACCAAGAGAGAGACTTACATCTGCTGCTGCTG 786
Qy 721 CGCTCAAGGCTTAACCTCCAAACAGCAAGGGAGATGAAGTTTGAAGTGAAGAACTTAATC 780
Db 787 CGCTCAAGGCTTAACCTCCAAACAGCAAGGGAGATGAAGTTTGAAGTGAAGAACTTAATC 846
Qy 781 CTACACAAGAGCTACAGCGCTGACAGCTTGTCTACACCAACGACATTCCTTGTGCTGAAG 840
Db 847 CTACACAAGAGCTACAGCGCTGACAGCTTGTCTACACCAACGACATTCCTTGTGCTGAAG 906
Qy 841 ATCCGTTCCAAAGAGGGCAGGTGTGCGACGCCATTCGCGAGCTATACAGACCATCTGCTG 900
Db 907 ATCCGTTCCAAAGAGGGCAGGTGTGCGACGCCATTCGCGAGCTATACAGACCATCTGCTG 966
Qy 901 CCTCGATGTAATAACGATCCCGAGTTTGGCAACAGTGTGAGATCACTGGCTTTGAGAAA 960
Db 967 CCTCGATGTAATAACGATCCCGAGTTTGGCAACAGTGTGAGATCACTGGCTTTGAGAAA 1026
Qy 961 GAGAAATTTCTACCGACTATCTTATTCGAGAGCAGCTGAAAAATGACTGTTGTGAAGCTGATT 1020
Db 1027 GAGAAATTTCTACCGACTATCTTATTCGAGAGCAGCTGAAAAATGACTGTTGTGAAGCTGATT 1086
Qy 1021 TCCCAACGGGAGGTGAGAGAGCCCACTTACAGGCTCTGAAAGTCAACCAAAATGCTG 1080
Db 1087 TCCCAACGGGAGGTGAGAGAGCCCACTTACAGGCTCTGAAAGTCAACCAAAATGCTG 1146
Qy 1081 TGTGCTGCTGACCCACAGTGTGAAGAACAGATTCTGTGCAAGGAGAGCTCAGGGGGAACCCCTC 1140

Db 1147 TGTGCTGCTGACCCACAGTGTGAAGAAACAGATTCTGTCCAGGAGAGACTCAGGGGGAACCCCTC 1206
Qy 1141 GTCTGTTCCTCCCAAGGCGCGCATGACTTTGATGTGAATTGTGAGCTGGGGCGTGTGAGT 1200
Db 1207 GTCTGTTCCTCCCAAGGCGCGCATGACTTTGATGTGAATTGTGAGCTGGGGCGTGTGAGT 1266
Qy 1201 GCCCTGAAGGACCAAGCGCTCTTACAGAGAGTCTCACACTTCTTAACCTGTGATCCGC 1260
Db 1267 GCCCTGAAGGACCAAGCGCTCTTACAGAGAGTCTCACACTTCTTAACCTGTGATCCGC 1326
Qy 1261 AGTCAACCAAGAGAGAGAAATGGCCCTGCGCTCTGA 1296
Db 1327 AGTCAACCAAGAGAGAGAAATGGCCCTGCGCTCTGA 1362

RESULT 7
US-10-313-986-123
; Sequence 123, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-123

Query Match 100.0%; Score 1296; DB 15; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGAGCCCTGCTGGGCGGCGCTCTCTCTGCGCTCTGCTGCGAGCGACTCCAAAGGC 60
Db 67 ATGAGAGCCCTGCTGGGCGGCGCTCTCTCTGCGCTCTGCTGCGAGCGACTCCAAAGGC 126
Qy 61 AGCAATGAACCTTCACTCAAGTTCATCGCAATGTGACTGTCTAAATVGAGAGAACATGTGTG 120
Db 127 AGCAATGAACCTTCACTCAAGTTCATCGCAATGTGACTGTCTAAATVGAGAGAACATGTGTG 186
Qy 121 TCCAAACAAGTACTTCTCCAAATTCATCTGTGTGCACTGTGCCCAAGAAATTCGAGGGCAG 180
Db 187 TCCAAACAAGTACTTCTCCAAATTCATCTGTGTGCACTGTGCCCAAGAAATTCGAGGGCAG 246
Qy 181 CACTGTGAATATAGTAAAGTCAAAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGGA 240
Db 247 CACTGTGAATATAGTAAAGTCAAAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGGA 306
Qy 241 AAGGCCAGCACTGACACCAATGGGCGGCGCTGCTGCTGGAACCTGTGCACTGTCTCTT 300
Db 307 AAGGCCAGCACTGACACCAATGGGCGGCGCTGCTGCTGGAACCTGTGCACTGTCTCTT 366
Qy 301 CAGCAACGTAACCACTGCGCCACAGATCTGATGCTCTTACGTGGGCGCTGGGGAACATTAAT 360
Db 367 CAGCAACGTAACCACTGCGCCACAGATCTGATGCTCTTACGTGGGCGCTGGGGAACATTAAT 426
Qy 361 TACTGCAGGAACCCAGCAACCCGAGGCGAACCTGTGTCTATGTGCAAGTGGGCTTAAAG 420
Db 427 TACTGCAGGAACCCAGCAACCCGAGGCGAACCTGTGTCTATGTGCAAGTGGGCTTAAAG 486
Qy 421 CCGCTGTCCAAAGTGCATGTGTGATGATCTGGCGCAATGGAAAAAGCCCTCTCTCTCT 480
Db 487 CCGCTGTCCAAAGTGCATGTGTGATGATCTGGCGCAATGGAAAAAGCCCTCTCTCTCT 546

Db 861 CTACCAAGACTACAGCGCTGACACGCTTGTCTACCAACAGACANTTGCCTTGAAG 920
QY 861 ATCCGTTCCAAAGAGAGGAGGTGTGTGCGAGCATTCGCCGACTATACAGACCATTCGCTG 900
Db 921 ATCCGTTCCAAAGAGAGGAGGTGTGTGCGAGCATTCGCCGACTATACAGACCATTCGCTG 980
QY 901 CCCTGATGTATTAACGATCCCAAGTTGGGCAAGCTGTGATGATCACTGGCTTTGGAAA 960
Db 961 CCCTGATGTATTAACGATCCCAAGTTGGGCAAGCTGTGATGATCACTGGCTTTGGAAA 1040
QY 961 GAGAAATTCACCGACTATCTCTATCCGAGCAGCTGAAAATGACTGTGTGTAAGCTATT 1020
Db 1041 GAGAAATTCACCGACTATCTCTATCCGAGCAGCTGAAAATGACTGTGTGTAAGCTATT 1100
QY 1021 TCCCAACGGGAGGTGTGACAGACCCCACTACTAGGCTGTGAAAGTCCACCAAAATGCTG 1080
Db 1101 TCCCAACGGGAGGTGTGACAGACCCCACTACTAGGCTGTGAAAGTCCACCAAAATGCTG 1160
QY 1081 TGTCTGTGTGACCCACAGTGGAAAACAGATTCTGTCCAGAGGAGACTCAGGGGGGACCCCTC 1140
Db 1161 TGTCTGTGTGACCCACAGTGGAAAACAGATTCTGTCCAGAGGAGACTCAGGGGGGACCCCTC 1220
QY 1141 GTCTGTTCCTCTCAGAGCGCGCATGACTTTGAATGTGTAAGCTGTGGGCGCTGTGATGT 1200
Db 1221 GTCTGTTCCTCTCAGAGCGCGCATGACTTTGAATGTGTAAGCTGTGGGCGCTGTGATGT 1280
QY 1201 GCCCTGAAGAGCAAGCGAGGTGTCTACAGAGTCTCAACTTCTTACCTTGAATCCG 1260
Db 1281 GCCCTGAAGAGCAAGCGAGGTGTCTACAGAGTCTCAACTTCTTACCTTGAATCCG 1340
QY 1261 AGTCACACCAAGAGAAATGCGCTGCGCTCTGA 1296
Db 1341 AGTCACACCAAGAGAAATGCGCTGCGCTCTGA 1376

RESULT 9
US-09-850-716A-122
; Sequence 122, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ. ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ. ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-850-716A-122

Query Match 99.8%; Score 1292.8; DB 9; Length 1475;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTGTGTGCGCGCCCTTCTGTGTGTCTGTGTGTGTGAGCACTTCCAAAGGC 60
Db 81 ATGAGAGCCCTGTGTGCGCGCCCTTCTGTGTGTCTGTGTGTGTGAGCACTTCCAAAGGC 140
QY 61 AGCAATGAATTCATCAAGTTCATGAACTGTAAGTGTCTTAATGAGAAACATGTGTG 120
Db 141 AGCAATGAATTCATCAAGTTCATGAACTGTAAGTGTCTTAATGAGAAACATGTGTG 200
QY 121 TCCAAACAAGTACTTCTTCCAAACATTCAGTGTGCAACTGCGCAAAGAAATTTGGAAGGCGAG 180
Db 201 TCCAAACAAGTACTTCTTCCAAACATTCAGTGTGCAACTGCGCAAAGAAATTTGGAAGGCGAG 260
QY 181 CACTGTGAATAGATTAAGTCAAAAACCTGTCTATGAGGGGAATGTCACTTTTACCGAAGA 240

Db 261 CACTGTGAATAGATTAAGTCAAAAACCTGTCTATGAGGGGAATGTCACTTTTACCGAAGA 320
QY 241 AAGGCCAGACATGACACCATTTGGCCCGGCTTCTGCTTCTGGAATCTTGGCACTGTCTT 300
Db 321 AAGGCCAGACATGACACCATTTGGCCCGGCTTCTGCTTCTGGAATCTTGGCACTGTCTT 380
QY 301 CAGCAAACTTACCATGACCCACAGATCTGATGTCTTTCAGTGGGCTGCGGGAACATAT 360
Db 381 CAGCAAACTTACCATGACCCACAGATCTGATGTCTTTCAGTGGGCTGCGGGAACATAT 440
QY 361 TACTGAGGAACCCAGACACACCGAGGAGACCTGTGTCTATGTGACAGTGGGCTTAAG 420
Db 441 TACTGAGGAACCCAGACACACCGAGGAGACCTGTGTCTATGTGACAGTGGGCTTAAG 500
QY 421 CCGCTTGTCTCAAGAGTGTGATGTGATCTGCGCAGATGGAATAAAAGCCCTCTCTCT 480
Db 501 CCGCTTGTCTCAAGAGTGTGATGTGATCTGCGCAGATGGAATAAAAGCCCTCTCTCT 560
QY 481 CCAAGAAATTAATAATTTCAGTGTGCGCAAAAGACTGTGAGGCGCCGCTTTAAGATTAT 540
Db 561 CCAAGAAATTAATAATTTCAGTGTGCGCAAAAGACTGTGAGGCGCCGCTTTAAGATTAT 620
QY 541 GGGGGGAAATTCACACCATGAGAACCGAGCCCTGTGTGTGGGCTTCTACAGAGGAC 600
Db 621 GGGGGGAAATTCACACCATGAGAACCGAGCCCTGTGTGTGGGCTTCTACAGAGGAC 680
QY 601 CCGGGGGGCTGTCTACCTAGTGTGTGAGAGGAGCCCTCATCAGCCCTTGTGTGATG 660
Db 681 CCGGGGGGCTGTCTACCTAGTGTGTGAGAGGAGCCCTCATCAGCCCTTGTGTGATG 740
QY 661 AGCGCACACACTGTCTTATTTATTTACCAAGAAAGAGACTATCTGTCTACCTGAGT 720
Db 741 AGCGCACACACTGTCTTATTTATTTACCAAGAAAGAGACTATCTGTCTACCTGAGT 800
QY 721 CGCTCAAGGCTTAACTCCAAACAGCGAAGGGAGATGAAGTTGAGTGGGAAAACCTATAT 780
Db 801 CGCTCAAGGCTTAACTCCAAACAGCGAAGGGAGATGAAGTTGAGTGGGAAAACCTATAT 860
QY 781 CTACCAAGAGACTACAGGCTGACAGCCTTGTCTACCAACAGACATTTGCTGTGAAG 920
Db 861 CTACCAAGAGACTACAGGCTGACAGCCTTGTCTACCAACAGACATTTGCTGTGAAG 920
QY 841 ATCCGTTCCAAAGAGGAGGTGTGTGCGAGCATTCGCCGACTATACAGACCATTCGCTG 900
Db 921 ATCCGTTCCAAAGAGGAGGTGTGTGCGAGCATTCGCCGACTATACAGACCATTCGCTG 980
QY 901 CCCTGATGTATTAACGATCCCAAGTTGGGCAAGCTGTGAGATCACTGGCTTTGAAA 960
Db 981 CCCTGATGTATTAACGATCCCAAGTTGGGCAAGCTGTGAGATCACTGGCTTTGAAA 1040
QY 961 GAGAAATTCACCGACTATCTCTATCCGAGCAGCTGAAAATGACTGTGTGTAAGCTATT 1020
Db 1041 GAGAAATTCACCGACTATCTCTATCCGAGCAGCTGAAAATGACTGTGTGTAAGCTATT 1100
QY 1021 TCCCAACGGGAGGTGTGACAGACCCCACTACTAGGCTGTGAAAGTCCACCAAAATGCTG 1080
Db 1101 TCCCAACGGGAGGTGTGACAGACCCCACTACTAGGCTGTGAAAGTCCACCAAAATGCTG 1160
QY 1081 TGTCTGTGTGACCCACAGTGGAAAACAGATTCTGTCCAGAGGAGACTCAGGGGGGACCCCTC 1140
Db 1161 TGTCTGTGTGACCCACAGTGGAAAACAGATTCTGTCCAGAGGAGACTCAGGGGGGACCCCTC 1220
QY 1141 GTCTGTTCCTCTCAGAGCGCGCATGACTTTGAATGTGTAAGCTGTGGGCGCTGTGATGT 1200
Db 1221 GTCTGTTCCTCTCAGAGCGCGCATGACTTTGAATGTGTAAGCTGTGGGCGCTGTGATGT 1280
QY 1201 GCCCTGAAGAGCAAGCGAGGTGTCTACAGAGTCTCAACTTCTTACCTTGAATCCG 1260
Db 1281 GCCCTGAAGAGCAAGCGAGGTGTCTACAGAGTCTCAACTTCTTACCTTGAATCCG 1340
QY 1261 AGTCACACCAAGAGAAATGCGCTGCGCTCTGA 1296

QY	1	ATGAGAGCCCTGCTGAGCGCGCGCTGCTTCCTGCGCGCTGCTGAGCGACTCCAAAGCC	60
Db	81	ATGAGAGCCCTGCTGAGCGCGCGCTGCTTCCTGCGCGCTGCTGAGCGACTCCAAAGCC	140
QY	61	AGCAATGAATCTTCATCAAGTTCCATCGAATCTGACTGTCTTAAATGAGAAACATGTGTG	120
Db	141	AGCAATGAATCTTCATCAAGTTCCATCGAATCTGACTGTCTTAAATGAGAAACATGTGTG	200
QY	121	TCGAACAAGTACTTCTCCAAACATTCACCTGTGTGCACTGTCCCAAGAAATTCGAGGGCAG	180
Db	201	TCGAACAAGTACTTCTCCAAACATTCACCTGTGTGCACTGTCCCAAGAAATTCGAGGGCAG	260
QY	181	CACGTGAAATAGATTAAGTCAAAAACCTGATAGAGGGAAATGTGCACTTTTACCAGGA	240
Db	261	CACGTGAAATAGATTAAGTCAAAAACCTGATAGAGGGAAATGTGCACTTTTACCAGGA	320
QY	241	AAGGCCGACACTGACACCATGTGGCGCGGCTGCTGCTGCCCTTGGAATCTGTGCCACTGTCTT	300
Db	321	AAGGCCGACACTGACACCATGTGGCGCGGCTGCTGCTGCCCTTGGAATCTGTGCCACTGTCTT	380
QY	301	CAGCAAAAGTACCAATGCGCCACAGATCTGATGCTCTTGAGCTGGGGCTGGGGAAACATAT	360
Db	381	CAGCAAAAGTACCAATGCGCCACAGATCTGATGCTCTTGAGCTGGGGCTGGGGAAACATAT	440
QY	361	TACTGCAAGAACCCAGACAACCCGAGGCGAACCTGTGTGCTATGTGCAAGTGTGGCTTAAG	420
Db	441	TACTGCAAGAACCCAGACAACCCGAGGCGAACCTGTGTGCTATGTGCAAGTGTGGCTTAAG	500
QY	421	CCGCTTGTCMAAGGTGCATGTGTGCAATGATGTGGGCAATGGAATAAAAGCCCTCTCTCT	480
Db	501	CCGCTTGTCMAAGGTGCATGTGTGCAATGATGTGGGCAATGGAATAAAAGCCCTCTCTCT	560
QY	481	CCGAAGAAATTAATAATTTCAGTGTGGCCAAAGAATCTGAGGCCCGCTTTAAGATTAT	540
Db	561	CCGAAGAAATTAATAATTTCAGTGTGGCCAAAGAATCTGAGGCCCGCTTTAAGATTAT	620
QY	541	GGGGGGAAGATTCAACACATCGAAGAACAGGCCCTGTGTTCCGGCATCTACAGAGGCAC	600
Db	621	GGGGGGAAGATTCAACACATCGAAGAACAGGCCCTGTGTTCCGGCATCTACAGAGGCAC	680
QY	601	CGGGGGGGCTCTGTCACTACCTGTGTGAGGAGAGCTCATCAAGCCCTTGCTGGGTGATC	660
Db	681	CGGGGGGGCTCTGTCACTACCTGTGTGAGGAGAGCTCATCAAGCCCTTGCTGGGTGATC	740
QY	661	AGGGCCACACCTGCTTCACTATGATTACCCAAAGAGAGGACTACATCGTCACTCTGGGT	720
Db	741	AGGGCCACACCTGCTTCACTATGATTACCCAAAGAGAGGACTACATCGTCACTCTGGGT	800
QY	721	CGCTCAAGGCTTAACCTCCAAACGCAAGGGGAGATGAAGTTGAGGTGGAAAACTTAATC	780
Db	801	CGCTCAAGGCTTAACCTCCAAACGCAAGGGGAGATGAAGTTGAGGTGGAAAACTTAATC	860
QY	781	CTACACAGAGCTACAGCGCTGACACGCTTGCTCACACACGACATTCGTTGCTGAAG	840
Db	861	CTACACAGAGCTACAGCGCTGACACGCTTGCTCACACACGACATTCGTTGCTGAAG	920
QY	841	ATCCGTTCCAAAGGAGGAGGTGTGTGGAGGCCATCCCGGACTATATACAGCACTGTGCTG	900
Db	921	ATCCGTTCCAAAGGAGGAGGTGTGTGGAGGCCATCCCGGACTATATACAGCACTGTGCTG	980
QY	901	CCCTGAGTATATACGATCCCAAGTTTGCAACAAGCTGTGAATCACTGAGCTTTGGAAAA	960
Db	981	CCCTGAGTATATACGATCCCAAGTTTGCAACAAGCTGTGAATCACTGAGCTTTGGAAAA	1040
QY	961	GAGAAATTCACGCACTATCTATCCGAGGAGCTGTAATAATGACTGTTGTGAAGCTGATT	1020
Db	1041	GAGAAATTCACGCACTATCTATCCGAGGAGCTGTAATAATGACTGTTGTGAAGCTGATT	1100
QY	1021	TCGCCACGGGAGTGTACAGAGCCCACTACTACGGCTCTGAAGTCAACCAACAAATGTGCTG	1080
Db	1101	TCGCCACGGGAGTGTACAGAGCCCACTACTACGGCTCTGAAGTCAACCAACAAATGTGCTG	1160
QY	1081	TGTGTGTGTGACCCCACTGTGAAAAACAGATTTCTTCCAGAGGAGACTCAAGGGGACCTCTC	1140

Db	1161	TGAGCTGCTGACCCCAAGTGGAAAAACAATTCTCTCCAGGGAGACTCAGGGGGACCCCTC	1220
Qy	1141	GTCTGTTCCCTCCAAAGCCCGCATGACTTTGACTGGATTGTGAGCTGGGGCCGTGATGT	1200
Db	1221	GTCTGTTCCCTCCAAAGCCCGCATGACTTTGACTGGATTGTGAGCTGGGGCCGTGATGT	1280
Qy	1201	GCCCTAAGACACAGCAGGGGGCTCACAAGAGATCTCACACTTCTTAACCTGATCCG	1260
Db	1281	GCCCTAAGACACAGCAGGGGGCTCACAAGAGATCTCACACTTCTTAACCTGATCCG	1340
Qy	1261	AGTCACACCAAGAAAGAAATGGCTGGCCCTCTGA	1296
Db	1341	AGTCACACCAAGAAAGAAATGGCTGGCCCTCTGA	1376
RESULT 12			
US-10-411-037-33			
; Sequence 33, Application US/10411037			
; Publication No. US20040043446A1			
; GENERAL INFORMATION:			
; APPLICANT: Neose Technologies, Inc.			
; APPLICANT: Defrees, Shawn			
; APPLICANT: Zopf, David			
; APPLICANT: Bayer, Robert			
; APPLICANT: Hakes, David			
; APPLICANT: Chen, Xi			
; APPLICANT: Bove, Caryn			
; TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF			
; FILE REFERENCE: 040853-01-5082			
; CURRENT APPLICATION NUMBER: US/10/411,037			
; CURRENT FILING DATE: 2003-04-09			
; PRIOR APPLICATION NUMBER: US 60/328,523			
; PRIOR FILING DATE: 2001-10-10			
; PRIOR APPLICATION NUMBER: US 60/344,652			
; PRIOR FILING DATE: 2001-10-19			
; PRIOR APPLICATION NUMBER: US 60/387,292			
; PRIOR FILING DATE: 2002-06-07			
; PRIOR APPLICATION NUMBER: US 60/391,777			
; PRIOR FILING DATE: 2002-06-25			
; PRIOR APPLICATION NUMBER: US 60/396,594			
; PRIOR FILING DATE: 2002-07-17			
; PRIOR APPLICATION NUMBER: US 60/404,249			
; PRIOR FILING DATE: 2002-08-16			
; PRIOR APPLICATION NUMBER: US 60/407,527			
; PRIOR FILING DATE: 2002-08-28			
; NUMBER OF SEQ ID NOS: 75			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 33			
; LENGTH: 1475			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-411-037-33			
Query Match 99.8%; Score 1292.8; DB 12; Length 1475;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	ATGAGACCCCTGCTGGGCGGCGCTGCTTCTGGTCTGGTGTGTGAGGACATCCAAAGGC	60
Db	81	ATGAGACCCCTGCTGGGCGGCGCTGCTTCTGGTCTGGTGTGTGAGGACATCCAAAGGC	140
Qy	61	AGCAATGAACCTTCATCAAGTTCATGAACTGTGTAATGAGAAATGTGTG	120
Db	141	AGCAATGAACCTTCATCAAGTTCATGAACTGTGTAATGAGAAATGTGTG	200
Qy	121	TCGAACAAGTACTTTCGCAACATTCACCTGTGTGCACTGCCCCAAGAAATTGGAGGGCAG	180
Db	201	TCGAACAAGTACTTTCGCAACATTCACCTGTGTGCACTGCCCCAAGAAATTGGAGGGCAG	260
Qy	181	CACGTGCAATATGATTAAGTCAAAAACCTGTGAGGGGAAATGGCACTTTTACCGAGA	240

Db 261 CACTGTGAATAGATAGTCAAAAACCTGCTATAGAGGAGTGTCACTTTTACCGAGGA 320
 Qy 241 AAGGCCAGCATGACATCCATGGCCGGCCCTGCTGCTGCACTCTGCACTGCTCTT 300
 Db 321 AAGGCCAGCATGACATCCATGGCCGGCCCTGCTGCTGCACTCTGCACTGCTCTT 380
 Qy 301 CAGCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Db 381 CAGCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 440
 Qy 361 TACTGAGAAACCCAGCAACCCGAGCGACCTGTGTCTATGTGCAAGTGGGCTTAAG 420
 Db 441 TACTGAGAAACCCAGCAACCCGAGCGACCTGTGTCTATGTGCAAGTGGGCTTAAG 500
 Qy 421 CGGCTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 Db 501 CGGCTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 560
 Qy 481 CCAGAAGATTAATTAATTTTCACTGTGGCCAAAGACTCTGAGGCCCGCTTTAAGATTAT 540
 Db 561 CCAGAAGATTAATTAATTTTCACTGTGGCCAAAGACTCTGAGGCCCGCTTTAAGATTAT 620
 Qy 541 GGGGAGAAATTCACCAATCGAAGAACAGCCCTGTGTGGCCATCTACAGAGGAC 600
 Db 621 GGGGAGAAATTCACCAATCGAAGAACAGCCCTGTGTGGCCATCTACAGAGGAC 680
 Qy 601 GGGGGGGGCTCTGTACCTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 Db 681 GGGGGGGGCTCTGTACCTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 740
 Qy 661 AAGGCCACACACTGCTTCACTTATGATTAACCAAGAGAGGAGCTACATCGCTACCTGGGT 720
 Db 741 AAGGCCACACACTGCTTCACTTATGATTAACCAAGAGAGGAGCTACATCGCTACCTGGGT 800
 Qy 721 CGCTCAAGGCTTAATCTCAACAGCAAGAGGAGATGAGTTTGTGTGTGTGTGTGTGT 780
 Db 801 CGCTCAAGGCTTAATCTCAACAGCAAGAGGAGATGAGTTTGTGTGTGTGTGTGTGT 860
 Qy 781 CTACACAAAGATCTACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
 Db 861 CTACACAAAGATCTACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 920
 Qy 841 ATCCGTTCCAAAGAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
 Db 921 ATCCGTTCCAAAGAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 980
 Qy 901 CCTCGATGTATACGATCCCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 Db 981 CCTCGATGTATACGATCCCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1040
 Qy 961 GAGGAATTTCAACGATCTATCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
 Db 1041 GAGGAATTTCAACGATCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1100
 Qy 1021 TCCCAACCGGAGAGT 1080
 Db 1101 TCCCAACCGGAGAGT 1160
 Qy 1081 TGT 1140
 Db 1161 TGT 1220
 Qy 1141 GTCTGT 1200
 Db 1221 GTCTGT 1280
 Qy 1201 GGCCTGTAAGAGCAAGCCAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
 Db 1281 GGCCTGTAAGAGCAAGCCAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1340
 Qy 1261 AGTCAACCAAG 1296
 Db 1341 AGTCAACCAAG 1376

RESULT 13
 US-10-117-982-122
 ; Sequence 122, Application US/10117982
 ; Publication No. US20030138438A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Foy, Teresa M.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedwick, Thomas S.
 ; APPLICANT: Carter, Darick
 ; APPLICANT: Matanabe, Yoshihiro
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Mericle, Barbara
 ; APPLICANT: Spies, Gregory A.
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Wang, Tonglong
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.455C18
 ; CURRENT APPLICATION NUMBER: US/10/117,982
 ; NUMBER OF SEQ ID NOS: 484
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 122
 ; LENGTH: 1475
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-117-982-122

Query Match 99.8%; Score 1292.8; DB 14; Length 1475;

Best Local Similarity 99.8%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 ATGAGAGCCCTGT 60
 Qy 81 ATGAGAGCCCTGT 140
 Db 61 AGCAATGAATCTTCATCAAGTTCATGAACTGTGATGTCTTAATGAGAGACATGTGTG 120
 Qy 141 AGCAATGAATCTTCATCAAGTTCATGAACTGTGATGTCTTAATGAGAGACATGTGTG 200
 Db 121 TCCAAAGATCTTCTTCAACATTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
 Qy 201 TCCAAAGATCTTCTTCAACATTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 260
 Db 181 CACTGTGAATAGATAGTCAAAAACCTGCTATGAGGGGAAATGTCACTTTTACCGAGGA 240
 Qy 261 CACTGTGAATAGATAGTCAAAAACCTGCTATGAGGGGAAATGTCACTTTTACCGAGGA 320
 Db 241 AAGGCCAGCATGACATCCATGGCCGGCCCTGCTGCTGTGCACTGTCTCTT 300
 Qy 321 AAGGCCAGCATGACATCCATGGCCGGCCCTGCTGCTGTGCACTGTCTCTT 380
 Db 301 CAGCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Qy 381 CAGCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 440
 Db 361 TACTGAGAAACCCAGCAACCCGAGCGACCTGTGTCTATGTGCAAGTGGGCTTAAG 420
 Qy 441 TACTGAGAAACCCAGCAACCCGAGCGACCTGTGTCTATGTGCAAGTGGGCTTAAG 500
 Db 421 CGGCTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 Qy 501 CGGCTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 560
 Db 481 CCAGAAGATTAATTAATTTTCACTGTGGCCAAAGACTCTGAGGCCCGCTTTAAGATTAT 540
 Qy 561 CCAGAAGATTAATTAATTTTCACTGTGGCCAAAGACTCTGAGGCCCGCTTTAAGATTAT 620
 Db 541 GGGGAGAAATTCACCAATCGAAGAACAGCCCTGTGTGGCCATCTACAGAGGAC 600

Db	622	GGGGGAGAAATTCAACACACATCCGAAACACAGCCCTGGTTTCGGCCATCTACAGAGACAC	680
QY	601	CGGGGGGGGCTCTGTCACTACCTAGTGTGTGAGAGGAGGCTCATCAGCCCTGTGGTGATC	660
Db	681	CGGGGGGGGCTGTGTCACTACCTAGTGTGTGAGGAGGCTCATCAGCCCTGTGGTGATC	740
QY	661	AGGCGACACACCTGCTTCATTGATTACCCAAABAAGAGAGCTACATCGCTACCTGGGT	720
Db	741	AGGCGACACACCTGCTTCATTGATTACCCAAABAAGAGAGCTACATCGCTACCTGGGT	800
QY	721	CGCTCAGGCTTAACTTCCAAACGCAAGGGAGATGAATTGAGGTGAAAACTTAATC	780
Db	801	CGCTCAGGCTTAACTTCCAAACGCAAGGGAGATGAATTGAGGTGAAAACTTAATC	860
QY	781	CTACACAAAGACATACAGCGCTGACAGCGTTGCTACACCAAGACATTCCTTGTGAAG	840
Db	861	CTACACAAAGACATACAGCGCTGACAGCGTTGCTACACCAAGACATTCCTTGTGAAG	920
QY	841	ATCCGTTCCAAAGAGGGAGGTGTGTGCAAGCCATCCCGACTATACAGACCATCTGGCTG	900
Db	921	ATCCGTTCCAAAGAGGGAGGTGTGTGCAAGCCATCCCGACTATACAGACCATCTGGCTG	980
QY	901	CCCTGAGTGTATATAGATGCCAGTTGGCACAAGTGTGAATCACTGGCTTTGGAAA	960
Db	981	CCCTGAGTGTATATAGATGCCAGTTGGCACAAGTGTGAATCACTGGCTTTGGAAA	1040
QY	961	GAGAAATTTCAACGACTATCTTATCCGAGCAGCTGAAATGACTGTTGTGAAGCTGATT	1020
Db	1041	GAGAAATTTCAACGACTATCTTATCCGAGCAGCTGAAATGACTGTTGTGAAGCTGATT	1100
QY	1021	TCCGACCGGGAGTGTACAGAGGCCCACTACTACGCTCTGAAGTCAACCAACAAATGCTG	1080
Db	1101	TCCGACCGGGAGTGTACAGAGGCCCACTACTACGCTCTGAAGTCAACCAACAAATGCTG	1160
QY	1081	TGTGTCTGTGACCCACAGTGTAAAAACAATTCCTCCAGAGGAGACTCAGGGGGACCCCTC	1140
Db	1161	TGTGTCTGTGACCCACAGTGTAAAAACAATTCCTCCAGAGGAGACTCAGGGGGACCCCTC	1220
QY	1141	GTCGTTCCTCCAGAGGCCGATATGACTTTGACTGGAATTTGTGAGCTGGGGCCGTGATGT	1200
Db	1221	GTCGTTCCTCCAGAGGCCGATATGACTTTGACTGGAATTTGTGAGCTGGGGCCGTGATGT	1280
QY	1201	GCCCTGAAGACAAAGCCAGGCGTCTACACGAGAGTCTCAACTTTACCCCTGGATCCGC	1266
Db	1281	GCCCTGAAGACAAAGCCAGGCGTCTACACGAGAGTCTCAACTTTACCCCTGGATCCGC	1340
QY	1261	AGTACACCAAGAGAAAGATAGCGCTGGCCCTCTGA	1296
Db	1341	AGTACACCAAGAGAAAGATAGCGCTGGCCCTCTGA	1376

```

RESULT 14 -US-10-101-510-159
; US-10-101-510-159
; Sequence 159, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WANG, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 159
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-101-510-159

```

Query Match	99.8%; Score 1292.8; DB 14; Length 1475;
-------------	--

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	ATGAGAGCCGCTGGGGGGGCGCCCTGCTCTCTGAGGCTCTGCTGAGAGACTCCAAAGGC	60
Db	ATGAGAGCCCTGCTGGGGGGCGCCCTGCTCTCTGAGGCTCTGCTGAGAGACTCCAAAGGC	140
QY	61 AGCAATGAACCTTCATCAAGTTCCATGCACTGTGACTGTCTTAAATGGAGAACTATGTG	120
Db	141 AGCAATGAACCTTCATCAAGTTCCATGCACTGTGACTGTCTTAAATGGAGAACTATGTG	200
QY	121 TTCAAACAAGTACTTCTCCAAACTTCACCTGTGTCAACTGCCAAAGAAATTCGGAGGGCAG	180
Db	201 TTCAAACAAGTACTTCTCCAAACTTCACCTGTGTCAACTGCCAAAGAAATTCGGAGGGCAG	260
QY	181 CACTGTGAAATAGATTAAGTCAAAAACCTGTCTATGAGGGGAAATGGTCACTTTTACCGAGA	240
Db	261 CACTGTGAAATAGATTAAGTCAAAAACCTGTCTATGAGGGGAAATGGTCACTTTTACCGAGA	320
QY	241 AAGGCCAGCACTGACACCAATGGGCGGCGCCCTGCTGCTGCGCTGTGAACTGTGCACCTGTCTT	300
Db	321 AAGGCCAGCACTGACACCAATGGGCGGCGCCCTGCTGCTGCGCTGTGAACTGTGCACCTGTCTT	380
QY	301 CAGCAAACTGATACCATCCCAAGATCTGATGTCTTTCAGCTGTGGGCTGTGGGAAACATPAT	360
Db	381 CAGCAAACTGATACCATCCCAAGATCTGATGTCTTTCAGCTGTGGGCTGTGGGAAACATPAT	440
QY	361 TACTGCAGGAACCCAGACAAACCGAGGCGACCCCTGTGTCTATGTGACAGTGGGCTTAAAG	420
Db	441 TACTGCAGGAACCCAGACAAACCGAGGCGACCCCTGTGTCTATGTGACAGTGGGCTTAAAG	500
QY	421 CCGCTTGTCCAAAGATGCAATGGTGTGATGATCTGCGCAGATGGAAAAAAGCCCTCTCTCCT	480
Db	501 CCGCTTGTCCAAAGATGCAATGGTGTGATGATCTGCGCAGATGGAAAAAAGCCCTCTCTCCT	560
QY	481 CCGAAGAAATTAATAATTTCAGTGTGGCCAAAGAATCTGAGGCCCGGCTTTAAGATTATT	540
Db	561 CCGAAGAAATTAATAATTTCAGTGTGGCCAAAGAATCTGAGGCCCGGCTTTAAGATTATT	620
QY	541 GGGGGAGAAATTCACCAACCATGAGAAACCAAGCCCTGTGTTCGGCCCATTTACAGAGAGCAC	600
Db	621 GGGGGAGAAATTCACCAACCATGAGAAACCAAGCCCTGTGTTCGGCCCATTTACAGAGAGCAC	680
QY	601 CGGGGGGGGCTCTGTCACTGCTGTGTGAGAGGACCTCATACAGCCCTTGTGGGTGATC	660
Db	681 CGGGGGGGGCTCTGTCACTGCTGTGTGAGAGGACCTCATACAGCCCTTGTGGGTGATC	740
QY	661 AAGGCCACACACTGTGCTTCAATTGATTAACCAAGAAAGAGAGACTATCTGCTTACCTGGGT	720
Db	741 AAGGCCACACACTGTGCTTCAATTGATTAACCAAGAAAGAGAGACTATCTGCTTACCTGGGT	800
QY	721 CGCTCAAGGCTTMACTCCACAACGCAAGGGGAGATGAATTTAGGTGTGAAAACTTAATC	780
Db	801 CGCTCAAGGCTTMACTCCACAACGCAAGGGGAGATGAATTTAGGTGTGAAAACTTAATC	860
QY	781 CTACACAAGAGATTAAGAGGCTGACAAGCTTGTCTACCAACAAGCAATTCGCTTGTGAAG	840
Db	861 CTACACAAGAGATTAAGAGGCTGACAAGCTTGTCTACCAACAAGCAATTCGCTTGTGAAG	920
QY	841 ATTCGTTCCAAGAGAGGAGGTGTGGCCAGCAATCCCGAGCTATACAGACCAATCTGCGCTG	900
Db	921 ATTCGTTCCAAGAGAGGAGGTGTGGCCAGCAATCCCGAGCTATACAGACCAATCTGCGCTG	980
QY	901 CCTCGATGTATTAACGATCCCAAGTTTGGCAACAGCTGTGAGATCACTGGCTTTGGAAAA	960
Db	981 CCTCGATGTATTAACGATCCCAAGTTTGGCAACAGCTGTGAGATCACTGGCTTTGGAAAA	1040
QY	961 GAGAAATTCACCGAATCTCTATCCGAGACACATGAATAATGATCTGTGTGAAGCTGATT	1020
Db	1041 GAGAAATTCACCGAATCTCTATCCGAGACACATGAATAATGATCTGTGTGAAGCTGATT	1100
QY	1021 TTCACAACGGAGATGTACAGACGCCCACTTACAGGCTCTGAAGTACCAACAAATATGCTG	1080

```

Db      1101 TCCACCGGGAGTGTACAGACCCCACTACTAGCGCTGTAGTACCAACCAAAATCTG 1160
Qy      1081 TGTGCTGTGACCCACAGTGGAAAAAGATTTCTGTCCAGAGGAGACTAGAGGGGACCCCTC 1140
Db      1161 TGTGCTGTGACCCACAGTGGAAAAAGATTTCTGTCCAGAGGAGACTAGAGGGGACCCCTC 1220
Qy      1141 GTCTGTTCCTCCAGAGCCGCGCATGACTTTGACTGAAATTTGAGAGCTGGGCGGTGAATG 1200
Db      1221 GTCTGTTCCTCCAGAGCCGCGCATGACTTTGACTGAAATTTGAGAGCTGGGCGGTGAATG 1280
Qy      1201 GCCCTGAAGAGACAGCCAGCGCTCTACAGAGAGTCTCACACTTTTACCTTGATCCG 1260
Db      1281 GCCCTGAAGAGACAGCCAGCGCTCTACAGAGAGTCTCACACTTTTACCTTGATCCG 1340
Qy      1261 AGTCACACCAAGAAAGAAATGGCTTGGCCCTCTGA 1296
Db      1341 AGTCACACCAAGAAAGAAATGGCTTGGCCCTCTGA 1376

RESULT 15
US-10-313-986-122
; Sequence 122, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Katanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-122

Query Match      99.8%; Score 1292.8; DB 15; Length 1475;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ATGAGAGCCCTGTGCGCGCGCTGTCTTCTGCTGCTGCTGAGCGACTCCAAAGCG 60
Db      81 ATGAGAGCCCTGTGCGCGCGCGCTGTCTTCTGCTGCTGCTGAGCGACTCCAAAGCG 140
Qy      61 AGCAATGAACCTTCATCAAGTTCATGAACTGTGACTGTCTAAATGAGAGAACATGTGTG 120
Db      141 AGCAATGAACCTTCATCAAGTTCATGAACTGTGACTGTCTAAATGAGAGAACATGTGTG 200
Qy      121 TCCAAAGATGACTTCCCAACATTCATCTGTGCACTGCCCAAGAAATTCGAGAGGCGAG 180
Db      201 TCCAAAGATGACTTCCCAACATTCATCTGTGCACTGCCCAAGAAATTCGAGAGGCGAG 260
Qy      181 CACTGTGAATAGATAAGTCAAAAACTGCTATGAGAGGAAATGTCATTTTACGAGGA 240
Db      261 CACTGTGAATAGATAAGTCAAAAACTGCTATGAGAGGAAATGTCATTTTACGAGGA 320
Qy      241 AAGGCGAGCACTGACACCAATGAGCGCGCGCTGCTGCTGCAACTCTGTGCACTGTCTT 300
Db      321 AAGGCGAGCACTGACACCAATGAGCGCGCGCGCTGCTGCTGCAACTCTGTGCACTGTCTT 380
Qy      301 CAGCAAAAGTAACATGCGCAAGATGATGATGCTTCAAGTGGCGCTGGGGGAAACATAAT 360
Db      381 CAGCAAAAGTAACATGCGCAAGATGATGATGCTTCAAGTGGCGCTGGGGGAAACATAAT 440
Qy      361 TACTGAGAGAACCCAGACAAACCGAGAGCGACCTGTGTGCTATGTGACAGTGGGCTTAAG 420
Db      441 TACTGAGAGAACCCAGACAAACCGAGAGCGACCTGTGTGCTATGTGACAGTGGGCTTAAG 500

```

```

Qy      421 CGGCTGTCCAAAGTGCATGNGCATGACTGTGCGAGATGGAAGAAAGCCCTCTCTCT 480
Db      501 CGGCTGTCCAAAGTGCATGNGCATGACTGTGCGAGATGGAAGAAAGCCCTCTCTCTCT 560
Qy      481 CCAGAAAGATTAATAATTTAGTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAAT 540
Db      561 CCAGAAAGATTAATAATTTAGTGTGCGCAAAAGACTCTGAGGCCCGCTTTAAGATTAAT 620
Qy      541 GGGGGAGAAATTCACCAACCATGAGAACAGCCCTGTGTTGGGCACTTACAGAGGAC 600
Db      621 GGGGGAGAAATTCACCAACCATGAGAACAGCCCTGTGTTGGGCACTTACAGAGGAC 680
Qy      601 GGGGGGGGCTGTCTGACCTGATGAGTGTGAGAGAGCTCTACAGCCCTGTGAGTATC 660
Db      681 GGGGGGGGCTGTCTGACCTGATGAGTGTGAGAGAGCTCTACAGCCCTGTGAGTATC 740
Qy      661 AGGCCACACACTGCTTCAATTAATTAACCAAAAGAGAGACTACATCGTCTACCTGGT 720
Db      741 AGGCCACACACTGCTTCAATTAATTAACCAAAAGAGAGACTACATCGTCTACCTGGT 800
Qy      721 CGCTCAAGGCTTAATCTCCAAACAGCAAGGGAGATGAAGTTGAGGTGAAAACTTAATC 780
Db      801 CGCTCAAGGCTTAATCTCCAAACAGCAAGGGAGATGAAGTTGAGGTGAAAACTTCATC 860
Qy      781 CTACAGAGAGACTACAGCGCTGACAGCTTGTCTACACACAAAGCAATGCTGCTGAAG 840
Db      861 CTACAGAGAGACTACAGCGCTGACAGCTTGTCTACACACAAAGCAATGCTGCTGAAG 920
Qy      841 ATCCGTTCCAAAGAGGAGAGTGTGCGAGCCATCCCGGACTTAAAGACATCTGCTG 900
Db      921 ATCCGTTCCAAAGAGGAGAGTGTGCGAGCCATCCCGGACTTAAAGACATCTGCTG 980
Qy      901 CCTCGATGTATTAACGATCCCAAGTTTGGCAAGCTGTGAGATCACTGCTTTGAAAA 960
Db      981 CCTCGATGTATTAACGATCCCAAGTTTGGCAAGCTGTGAGATCACTGCTTTGAAAA 1040
Qy      961 GAGAAATTCACCAAGACTATCTTATCCGAGAGAGCTGAAAAATGACTTTGAGAGCTGAT 1020
Db      1041 GAGAAATTCACCAAGACTATCTTATCCGAGAGAGCTGAAAAATGACTTTGAGAGCTGAT 1100
Qy      1021 TCCCAACGGGAGTGTACAGAGCCCACTACTAGCGCTGTGAAGTCAACCAAAATGCTG 1080
Db      1101 TCCCAACGGGAGTGTACAGAGCCCACTACTAGCGCTGTGAAGTCAACCAAAATGCTG 1160
Qy      1081 TGTGCTGTGACCCACAGTGGAAAAAGATTTCTGTCCAGAGGAGACTCAAGGGGACCCCTC 1140
Db      1161 TGTGCTGTGACCCACAGTGGAAAAAGATTTCTGTCCAGAGGAGACTCAAGGGGACCCCTC 1220
Qy      1141 GTCTGTTCCTCCAGAGCCGCGCATGACTTTGACTGAAATTTGAGAGCTGGGCGGTGAATG 1200
Db      1221 GTCTGTTCCTCCAGAGCCGCGCATGACTTTGACTGAAATTTGAGAGCTGGGCGGTGAATG 1280
Qy      1201 GCCCTGAAGAGACAGCCAGCGCTCTACAGAGAGTCTCACACTTTTACCTTGATCCG 1260
Db      1281 GCCCTGAAGAGACAGCCAGCGCTCTACAGAGAGTCTCACACTTTTACCTTGATCCG 1340
Qy      1261 AGTCACACCAAGAAAGAAATGGCTTGGCCCTCTGA 1296
Db      1341 AGTCACACCAAGAAAGAAATGGCTTGGCCCTCTGA 1376

```

Search completed: March 21, 2004, 21:00:58
Job time : 522 secs

THIS PAGE BLANK (USPTO)

FEATURES

Location/Qualifiers
 1. 870
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6292200"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_43"
 /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library." }

ORIGIN

Query Match 63.5%; Score 822.6; DB 13; Length 870;
 Best Local Similarity 98.7%; Pred. No. 1.9e-213;
 Matches 850; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

75 TCAAGTTCATCGAAGCTGTGACTGTCTAAATGAGAGAACTGTGTCTCAACAAGTACTT 134
 1 TCAAGTTCATCGAAGCTGTGACTGTCTAAATGAGAGAACTGTGTCTCAACAAGTACTT 60

135 CTCCACATTCACCTGGTGGCAATCCGCCAAGAAATTTGGAGGGCAGCAGCTGTGAATAGA 194
 61 CTCCACATTCACCTGGTGGCAATCCGCCAAGAAATTTGGAGGGCAGCAGCTGTGAATAGA 120

195 TAACTCAAAAACCTGTATGAGGGGAAATGGTCACTTTTACCGAGAAAGCCAGCACTGA 254
 121 TAACTCAAAAACCTGTATGAGGGGAAATGGTCACTTTTACCGAGAAAGCCAGCACTGA 180

255 CACATGGGCGGCGCTGCTGCTGCTGAACTCTGCACTGTCTTCAAGCAACGTAACA 314
 181 CACATGGGCGGCGCTGCTGCTGCTGAACTCTGCACTGTCTTCAAGCAACGTAACA 240

315 TGCCACAGATCTGATGCTCTTCAAGTGGGCTGGGGAAACATATTAATTCAGAGAAACC 374
 241 TGCCACAGATCTGATGCTCTTCAAGTGGGCTGGGGAAACATATTAATTCAGAGAAACC 300

375 AGACAAACCGAGGCGGACCTGTGCTATGTGCAAGTGGGCTTAAAGCCGCTTGTCCAGA 434
 301 AGACAAACCGAGGCGGACCTGTGCTATGTGCAAGTGGGCTTAAAGCCGCTTGTCCAGA 360

435 GTGCATGGTGCAGACCTGGCGCAGATGAAAGAAAGCCCTCTCTCTCCAGAGATTTAA 494
 361 GTGCATGGTGCAGACCTGGCGCAGATGAAAGAAAGCCCTCTCTCTCCAGAGATTTAA 420

495 ATTTCAGTGTGGCAAAAGACTCTGAGGCCCCGCTTTAAGATTATTTGGGGAGAAATTC 554
 421 ATTTCAGTGTGGCAAAAGACTCTGAGGCCCCGCTTTAAGATTATTTGGGGAGAAATTC 480

555 CACCATTCGAGAACCAAGCCCTGTGTTGGGCTATCTACAGAGGACACGGGGGGCTCTGT 614
 481 CACCATTCGAGAACCAAGCCCTGTGTTGGGCTATCTACAGAGGACACGGGGGGCTCTGT 540

615 CACCTACGTGTGGAGGCGAGCTCATACGCCCTTCTGGGTGATATAGAGCCACACACTG 674
 541 CACCTACGTGTGGAGGCGAGCTCATACGCCCTTCTGGGTGATATAGAGCCACACACTG 600

675 CTTCATTGATTATCCCAAAGAGAGACTACATCGTCTACCTGGTGGCTCAAGGCTTAA 734
 601 CTTCATTGATTATCCCAAAGAGAGACTACATCGTCTACCTGGTGGCTCAAGGCTTAA 660

735 CTTCACAGGCAAGGGGAGATGAAGTTTGAAGTGGAAACCTTAATCTTACACAGAGACTA 794
 661 CTTCACAGGCAAGGGGAGATGAAGTTTGAAGTGGAAACCTTAATCTTACACAGAGACTA 720

795 CAGCGCTGACAGGCTTGTCAACACAGACATTTGCTTGAAGATCCGTT -CCAAG 853

Db 721 CAGCGCTGACAGGCTTGTCTACACACATGACATGCTTGTGAAGATCCGTTCCCAAG 780

Qy 854 AGGCGAGGTTGGGAGGACATCCCGGACTATACAGACATCTGCGCTG -CCCTCGATGAT 912

Db 781 AGGCGAGGTTGGGAGGACATCCCGGACTATACAGACATCTGCGCTTGTGAAGATCC 840

Qy 913 AACGATCCCGAGTTGGCACA 933

Db 841 AACGATCCCGAGTTGGGCA 861

RESULT 2

CA489682 882 bp mRNA linear EST 14-NOV-2002
 LOCUS CA489682
 DEFINITION AGENCOURT 10810373 MAFcl Homo sapiens cDNA clone IMAGE:6722162 5', mRNA sequence.
 ACCESSION CA489682
 VERSION CA489682.1 GI:24952473
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE

AUTHORS NIH-MGC http://mgc.mci.nih.gov/
 TITLE NIH-MGC cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Kristi A. Eglund, Ira Pastan
 cDNA Library Preparation: Invitrogen Corp
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
 Plate: LLM14284 row: h column: 02
 High quality sequence start: 22
 High quality sequence stop: 640.
 Location/Qualifiers
 1. 882
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6722162"
 /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, INCaP"
 /lab_host="EMDH10B"
 /clone_id="MAPCL"
 /note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dt. Average insert size: 1800 bp. Library amplification: 26,000 fold. Bunkook Lee & Ira Pastan. Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

FEATURES

source

Query Match 62.5%; Score 809.6; DB 14; Length 882;
 Best Local Similarity 96.6%; Pred. No. 7e-210;
 Matches 847; Conservative 0; Mismatches 27; Indels 3; Gaps 2;

127 AAGTACTTCCCAACATTCACTGCTGCACTGCCCAAGAAATTTGGAGGGCGGCACTGT 186
 1 AAGGCTTGTAACNNCGGTCCGAATTCNCGGAAATTCGAGAGGCGGCACTGT 60

Db 1 AAGGCTTGTAACNNCGGTCCGAATTCNCGGAAATTCGAGAGGCGGCACTGT 60

Qy 187 GAAATAGATTAAGTCAAAAACCTGCTATGAGGGGAATGTCTATTACCGAGGAAAGGCC 246
 61 GAAATAGATTAAGTCAAAAACCTGCTATGAGGGGAATGTCTATTACCGAGGAAAGGCC 120

Db 247 AGCAGTACACATGGGCGGCGCTGCTGCTGGAATCTTGCACATGCTCTTCAGCAA 306

```

Db      121 AGCACTGACACCATGAGCGCCCTGCTGCTGCTGCACTGCTTCAGCAA 180
QY      307 ACCTACCATGCCACAGATCTGATGCTCTTCACTGGGCTGGGAAACATTAATTCTGC 366
Db      181 AGGTACCAATGCCACAGATCTGATGCTCTTCACTGGGCTGGGAAACATTAATTCTGC 240
QY      367 AGGAACCCAGACACACGAGGCGACCCCTGGTGTATGTGAGGTGGGCTTAAAGCGCTT 426
Db      241 AGGAACCCAGACACACGAGGCGACCCCTGGTGTATGTGAGGTGGGCTTAAAGCGCTT 300
QY      427 GTTCAAGAGTGCATGTGTGATGATCTGCGCAGATGAGAAAAAGCCCTCTCTCTCCAGAA 486
Db      301 GTTCAAGAGTGCATGTGTGATGATCTGCGCAGATGAGAAAAAGCCCTCTCTCTCCAGAA 360
QY      487 GAATTAATAATTAGTGTGCGCCAAAAGATCTGTAGAGCCCGCTTTAAGATTATTTGGGGGA 546
Db      361 GAATTAATAATTAGTGTGCGCCAAAAGATCTGTAGAGCCCGCTTTAAGATTATTTGGGGGA 420
QY      547 GAATTCACCAACCATGAGAACAGCCCTGGTGTGCGGCGCATCTACAGAGAGGACCGGGGG 606
Db      421 GAATTCACCAACCATGAGAACAGCCCTGGTGTGCGGCGCATCTACAGAGAGGACCGGGGG 480
QY      607 GGCTCTGTCACTCACTGCTGTGTGAGGACAGCCCTTCATCAGCCCTTGTGTGATCAGCGCC 666
Db      481 GGCTCTGTCACTCACTGCTGTGTGAGGACAGCCCTTCATCAGCCCTTGTGTGATCAGCGCC 540
QY      667 ACACACTGCTTCTTCTTATTTATCCCAAGAAAGAGAGACTATCATGCTTACTGCGGTGCTCA 726
Db      541 ACACACTGCTTCTTCTTATTTATCCCAAGAAAGAGAGACTATCATGCTTACTGCGGTGCTCA 600
QY      727 AGGCTTAACCTCCAAACGAGAGGAGAGATGAGATTTGAGGTGAGAAAAACCTATCTTACAC 786
Db      601 AGGCTTAACCTCCAAACGAGAGGAGAGATGAGATTTGAGGTGAGAAAAACCTATCTTACAC 660
QY      787 AAGGACTACAGCGCTGACACAGCTTGTCTCACACAGACATTCCTTGTGAGATCCGT 846
Db      661 AAGGACTACAGCGCTGACACAGCTTGTCTCACACAGACATTCCTTGTGAGATCCGT 720
QY      847 TCCAGAGAGGCGAGGTGTGCGCAGCCATCCCGGACTATACAGACATTCGCTGCGCTCG 906
Db      721 TCCAGAGAGGCGAGGTGTGCGCAGCCATCCCGGACTATACAGACATTCGCTGCGCTCG 780
QY      907 ATGTATTAAGATGCCAGATTTGGGACAGAGT-GTGAATCAGCTG--CTTTGAGAAAAAG 963
Db      781 ATGTATTAAGATGCCAGATTTGGGACAGAGTGTGAATCAGCTGAGCTTTGAGAAAAAG 840
QY      964 AATTCTACCGACTATCTCTATCCGAGAGAGCTGAAAA 1000
Db      841 AATTCTACCGACTATCTCTATCCGAGAGAGCTGAAAA 877

RESULT 3
CF132029      820 bp      mRNA      linear      EST 05-AUG-2003
LOCUS      UI-HF-P00-aws-o-20-0-UI_r1.NIH.MGC_215 Homo sapiens cDNA clone
DEFINITION      IMAGE:30553987.5', mRNA sequence.
ACCESSION      CF132029
VERSION      CF132029.1 GI:33214878
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      Bonaldo,M.F., Lennon,G. and Soares,M.B.
AUTHORS      Normalisation and subtraction: two approaches to facilitate gene
TITLE      discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
PUBMED      8889548
COMMENT      Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa

```

```

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: Bento-Soares@iowa.edu
Tissue Procurement: Mary Hendrix
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.
Location/Qualifiers
1. 820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30553987"
/lab host="DH10B (T1 phage resistant)"
/clone_1ib="NIH MGC 215"
/notes="Organ: Lung; Vector: pYX-Asc; Site:1: EcoR I;
Site:2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

```

ORIGIN

```

Query Match      61.9%; Score 802.2; DB 14; Length 820;
Best Local Similarity 99.3%; Pred. No. 7e-208;
Matches 815; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

```

QY      218 GGAATGTCACCTTTACCGAGAAAGCCAGACATGACACCATGAGCCCGGCTGCTGC 277
Db      1 GGAATGTCACCTTTACCGAGAAAGCCAGACATGACACCATGAGCCCGGCTGCTGC 60
QY      278 CCTGAACCTGCGCACTGTCTTCTCAGCAAAAGTACCATGCGCCACAGATCTGATCTTTC 337
Db      61 CCTGAACCTGCGCACTGTCTTCTCAGCAAAAGTACCATGCGCCACAGATCTGATCTTTC 120
QY      338 AGCTGGGCTGGGAAACATTAATTACTGACAGAAACCCAGACACCGAGGCGACCTGTGT 397
Db      121 AGCTGGGCTGGGAAACATTAATTACTGACAGAAACCCAGACACCGAGGCGACCTGTGT 180
QY      398 GCTATGTGACAGTGGGCGCTTAAAGCCGCTTGTCCAAAGTGCATGTGATGCGGAG 457
Db      181 GCTATGTGACAGTGGGCGCTTAAAGCCGCTTGTCCAAAGTGCATGTGATGCGGAG 240
QY      458 ATGGAATAAAAGCCCTCTCTCTCCAGAAATTAATAATTAGTGTGCGCAAAAGACTC 517
Db      241 ATGGAATAAAAGCCCTCTCTCTCCAGAAATTAATAATTAGTGTGCGCAAAAGACTC 300
QY      518 TGAAGCCCGCTTTAAGATTATTTGGGGAGAAATTCACCAACATGAGAAACAGCCCTGT 577
Db      301 TGAAGCCCGCTTTAAGATTATTTGGGGAGAAATTCACCAACATGAGAAACAGCCCTGT 360
QY      578 TTGCGGCACTTACAGAGAGACACCGGGGGGCTCTGTCACTTACGCTGTGTGAGGAGCC 637
Db      361 TTGCGGCACTTACAGAGAGACACCGGGGGGCTCTGTCACTTACGCTGTGTGAGGAGCC 420
QY      638 TATACAGCCCTTGTGATGATCAGCGCCACACATGCTTCAATTGATTAACCAAAAGAG 697
Db      421 TATACAGCCCTTGTGATGATCAGCGCCACACATGCTTCAATTGATTAACCAAAAGAG 480
QY      698 AGGACTACATCGTCTTACCTGGGTGCTCAAGGCTTAATCTCAACAGCAAGGAGATGA 757
Db      481 AGGACTACATCGTCTTACCTGGGTGCTCAAGGCTTAATCTCAACAGCAAGGAGATGA 540

```

QY	758	AGTTGAGSTGGAAAACTTAATCCATACACAAAGACTACACGGCTATACAGCTTGGTCAC	817
Db	541	AGTTGAGSTGGAAAACTCATCTTACACAAAGACTACACGGCTATACAGCTTGGTCAC	600
QY	818	ACAACGACATTGGCTTGGTGAAGATCCGTTCCAGAGGGGAGGTGGCGAGCCATCCC	877
Db	601	ACAACGACATTGGCTTGGTGAAGATCCGTTCCAGAGGGGAGGTGGCGAGCCATCCC	660
QY	878	GGACATATACAAACCATCTGCGCCCTCGATAGTATAAGATCCCAAGTTGGCAACAGT	937
Db	661	GGACATATACAAACCATCTGCGCCCTCGATAGTATAAGATCCCAAGTTGGCAACAGT	720
QY	938	GTGAGATCACTGGCTTTGGAAAAAGAAATTCACGACTATCTATCCGAGACAGCTGA	997
Db	721	GTGAGATCACTGGCTTTGGAAAAAG-ATTCTACCGACTATCTATCCGAGACAGCTGA	779
QY	998	AAATGACTGTGTGAAGCTGATTTCCACACGGGAGATGTCAG 1038	
Db	780	NAATGACTGTGTGAAGCTGATTTCCACACGGGAGATGTCAG 820	

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 803)		NIH-MGC http://mgc.nci.nih.gov/ .		
		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.			

11556 Procurement: James Cleaver, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: LNLN0637 row: p column: 09
 High quality sequence stop: 784.

```

FEATURES
Source
Location/Qualifiers
1..803
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4779896"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1kb="NCI CGAP Skn3"
/name="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."

```

Query Match	60.5%	Score 783.6,	DB 12,	Length 803,
Best Local Similarity	99.4%	Pred. No. 8.5e-203,		
Matches 797; Conservative	0;	Mismatches 4;	Indels 1;	Gaps 1;

QY 224 GCACACTTTTACCGAGAAAGGCCAGCACTGACACCATGGCCGGCCCTGCTGCTTGA 283

Pb 1 GTCTACTTTTACCGAGAAAGGCCAGCACTGACACCATGGCCGGCCCTGCTGCTTGA 60

QY 284 ACTCTGGCACTGTCTCTTCAGCAACGTACCATCCTCCACAGATCTGTATCTCTTCAGCTGG 343

db 61 ACTCTGGCACTGTCTCTTCAGCAACGTACCATCCTCCACAGATCTGTATCTCTTCAGCTGG 120

QY	344	GCCTGGGGAAACATATATTACATGCGAAACCAGCAACCCGAGGCGACCTCGGTGCTATG	403
Db	121	GCTTGGGAAACATATATTACTGCGAAGAACCGACCAACCGAGGCGACCTCGGTGCTATG	180
QY	404	TGCAGGTGGGCTTAAAGCCGTTGTGCCAAGTGTCATGTGTGCATGACTGCGCAGATGGAA	463
Db	181	TGCAGGTGGGCTTAAAGCTGTGTGCCAAGTGTCATGTGTGCATGACTGCGCAGATGGAA	240
QY	484	AAAAGCCCTCCTCTCCCTCCAGAAATAATTAATTTTCAGTGTGGCCAAAACATCTGAGGC	523
Db	241	AAAAGCCCTCCTCTCTCCAGAAATAATTAATTTTCAGTGTGGCCAAAACATCTGAGGC	300
QY	524	CCCGCTTTAAGATTATTGGGGAGAAATTCACACATCGAGAAACAGCCCTGGTTTGGG	583
Db	301	CCCGCTTTAAGATTATTGGGGAGAAATTCACACACATCGAGAACAGCCCTGGTTTGGG	360
QY	584	CCATCTACAGAGGCAACCGGGGGGCTCTGTCACTACGTGTGTGAGGCGAGCCTCATCA	643
Db	361	CCATCTACAGAGGCAACCGGGGGGCTCTGTCACTACGTGTGTGAGGCGAGCCTCATCA	420
QY	644	GCCCTTGCTGGGTGATCAGCGCCACACACTGCTTCATTGATTAACCCAAAGAGAGACT	703
Db	421	GCCCTTGCTGGGTGATCAGCGCCACACACTGCTTCATTGATTAACCCAAAGAGAGACT	480
QY	704	ACATGCTCTACTTGGGTGCTCAAGGCTTAACTCCAAACAGCAAGGGGAGATGAAGTTTG	763
Db	481	ACATGCTCTACTTGGGTGCTCAAGGCTTAACTCCAAACAGCAAGGGGAGATGAAGTTTG	540
QY	764	AGGTGAAAAACCTAATCTTACACAAGGACTACAGCGCTGACAGCTTGCTACACAAGG	823
Db	541	AGGTGAAAAACCTAATCTTACACAAGGACTACAGCGCTGACAGCTTGCTACACAAGG	600
QY	824	ACATTGCCTTGCTGAAGATCCGTTCCAAAGGAGGCAAGTGTGCGACCATCCCGACTA	883
Db	601	ACATTGCCTTGCTGAAGATCCGTTCCAAAGGAGGCAAGTGTGCGACCATCCCGACTA	660
QY	884	TACAGACCATCTGCTGCGCCCTGCATGTATTAAGATATCCCAAG-TTTGGCACAGCTGTAG	942
Db	661	TACAGACCATCTGCTGCGCCCTGCATGTATTAAGATATCCCAAGTTTGGCACAGCTGTAG	720
QY	943	ATCACTGGCTTTGGAAAAAGAAATTCACGCACTATCTCATCCGAGACGCTGAATAATG	1002
Db	721	ATCACTGGCTTTGGAAAAAGAAATTCACGCACTATCTCATCCGAGACGCTGAATAATG	780
QY	1003	ACTGTTGTGAAGCTGATTTCCC	1024
Db	781	ACTGTTGTGAAGCTGATTTTCC	802

RESULT	5
LOCUS	BX360462
DEFINITION	BX360462 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSDD1070YD01 5-PRIME, mRNA sequence.
ACCESSION	BX360462
VERSION	BX360462.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1200)	Li, W. B., Gruber, C., Jesssee, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished (2001)	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr , Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 5613.f For more information about this cluster, see

http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D1070CB010P1&cluster=5613.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1070CB010P1.
 Location/Qualifiers

FEATURES

source

1.1200
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1070D01"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="Left strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 60.4%; Score 783.2; DB 13; Length 1200;
 Best Local Similarity 96.4%; Pred. No. 1.4e-202;
 Matches 812; Conservative 15; Mismatches 8; Indels 3; Gaps 3;

1 ATGAGAGCCCTGCTGGCGCCCTGCTCTCTGCGTCTGCTGCGAGCACTCCAAAGGC 60
 169 ATGAGAGCCCTGCTGGCGCCCTGCTCTCTGCGTCTGCTGCGAGCACTCCAAAGGC 228
 61 AGCAATGAATTCATCAAGTTCATCGAAGTCTGCTGCTTAATGAGAGAACTGTGTG 120
 229 AGCAATGAATTCATCAAGTTCATCGAAGTCTGCTGCTTAATGAGAGAACTGTGTG 288
 121 TCCAAAGAGTACTTCCCAATTCCTGCTGCAAGTCCCAAGAAATTCGAGAGGCAG 180
 289 TCCAAAGAGTACTTCCCAATTCCTGCTGCAAGTCCCAAGAAATTCGAGAGGCAG 348
 181 CACTGTGAATATAGTAAAGTCAAAACCTGCTATGAGGGGAATGCTCACTTTACCGAGA 240
 349 CACTGTGAATATAGTAAAGTCAAAACCTGCTATGAGGGGAATGCTCACTTTACCGAGA 408
 241 AAGGCGAGCACTGACACCATGAGGCGCGCCCTGCTGCTGCTGAACTGTGCACTGTCTT 300
 409 AAGGCGAGCACTGACACCATGAGGCGCGCCCTGCTGCTGCTGAACTGTGCACTGTCTT 468
 301 CAGCAATGATACATATCCCAAGATCTGATGCTCTTCACTGAGGCGCTGGGAAACATTAAT 360
 469 CAGCAATGATACATATCCCAAGATCTGATGCTCTTCACTGAGGCGCTGGGAAACATTAAT 528
 361 TACTGAGGAACCCAGAACCCGAGGCGAGCCTGCTGCTATGAGGAGTGGGCTTAAG 420
 529 TACTGAGGAACCCAGAACCCGAGGCGAGCCTGCTGCTATGAGGAGTGGGCTTAAG 587
 421 CCGCTTGTCCAGAGTGCATGATGATGCTGCGCAAGTGAAGAAAGCCCTCTCTCTCT 480
 588 CCGCTTGTCCAGAGTGCATGATGATGCTGCGCAAGTGAAGAAAGCCCTCTCTCTCTCT 647
 481 CCGAAGATTAATTTTCAGTGTGGCAAAAGACTCTGAGGCGCCGCTTTAAGATTATT 540
 648 CCGAAGATTAATTTTCAGTGTGGCAAAAGACTCTGAGGCGCCGCTTTAAGATTATT 707
 541 GGGGGAGAAATTCACCAATCGAGAACCAAGCCCTGGTTTGGGCAATCTACAGAGGAC 600
 708 GGGGGAGAAATTCACCAATCGAGAACCAAGCCCTGGTTTGGGCAATCTACAGAGGAC 767
 601 CCGGGGGGCTCTGTCACTGCTGTGAGGAGGCAAGCTCATGAGCCCTTGTGCTGATC 660
 768 CCGGGGGGCTCTGTCACTGCTGTGAGGAGGCAAGCTCATGAGCCCTTGTGCTGATC 827
 661 AGGCGCAACACCTGCTTCAATGATTAACCAAGAGAGAGACATACATGCTTACCTGGGT 720
 828 AGGCGCAACACCTGCTTCAATGATTAACCAAGAGAGAGACATACATGCTTACCTGGGT 887
 721 CGCTCAAGGCTTAATTCACCAAGAGAGAGAGATGAGTTGAGGTGAGAAACCTTAATC 780

Db 888 CGCTCAAGGCTTAATTCACCAAGAGAGAGAGATGAGTTGAGGTG-AAAACCTCATC 946
 Qy 781 CTACACAGAGACTACAGCGCTGACAGCGTGTCTTCAACACAGCAATTCCTTCTCTA 838
 947 MTACAAAGAGACTACAGCGSTGA-AGAGTGTCTCAAAAATGACATTCGCTTGTGA 1003

RESULT 6
 LOCUS BU190752 1003 bp mRNA linear EST 04-SEP-2002
 DEFINITION AGENCOURT_6925417 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5952454
 5', mRNA sequence.
 ACCESSION BU190752
 VERSION BU190752.1 GI:22704736
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1003)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC

REFERENCE
 AUTHORS NIH-MGC
 TITLE NIH-MGC
 JOURNAL
 COMMENT

CDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LHCW2140 row: h column: 23
 High quality sequence stop: 608.

FEATURES

source

1.1003
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5952454"
 /tissue_type="ductal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 110"
 /note="Organ: pancreas; Vector: pOT87; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 60.3%; Score 781; DB 13; Length 1003;
 Best Local Similarity 98.4%; Pred. No. 5e-202;
 Matches 809; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

1 ATGAGAGCCCTGCTGGCGCCCTGCTCTCTGCGTCTGCTGCGAGCACTCCAAAGGC 60
 73 ATGAGAGCCCTGCTGGCGCCCTGCTCTCTGCGTCTGCTGCGAGCACTCCAAAGGC 132
 61 AGCAATGAATTCATCAAGTTCATCGAAGTCTGCTGCTTAATGAGAGAACTGTGTG 120
 133 AGCAATGAATTCATCAAGTTCATCGAAGTCTGCTGCTTAATGAGAGAACTGTGTG 192
 121 TCCAAAGAGTACTTCCCAATTCCTGCTGCAAGTCCCAAGAAATTCGAGAGGCAG 180
 193 TCCAAAGAGTACTTCCCAATTCCTGCTGCAAGTCCCAAGAAATTCGAGAGGCAG 252
 181 CACTGTGAATATAGTAAAGTCAAAACCTGCTATGAGGGGAATGCTCACTTTACGAGGA 240
 253 CACTGTGAATATAGTAAAGTCAAAACCTGCTATGAGGGGAATGCTCACTTTACGAGGA 312

OY		241	AAGGCGACAGCATGCAACACATGGGCCCGGACCCTGCCTGCCTCCGGAACCTCAGACTGTCCTT	300
Db		313	AAGGCGACAGCATGCAACACATGGGCCCGGACCCTGCCTGCCTCCGGAACCTCAGACTGTCCTT	372
OY		301	CAGCAAAAGCTTACCATTGCCAACAAGATCTGATGTCTTTCACTGCGGCTTGCGAAAACATAAT	360
Db		373	CAGCAAAAGCTTACCATTGCCAACAAGATCTGATGTCTTTCACTGCGGCTTGCGAAAACATAAT	432
OY		361	TACTGCAGAACCCCAACAACCGGAGCGCACCCCTGGTGTATATGTGAGAGTGGGCCCTTAAG	420
Db		433	TACTGCAGAACCCCAACAACCGGAGCGCACCCCTGGTGTATATGTGAGAGTGGGCCCTTAAG	492
OY		421	CCGCTTGTCCAAGAAGTGATGGTGCA TGA CTGCGCAGATGAGAAAAAAGCCCTTCCTCCT	480
Db		493	CCGCTTGTCCAAGAAGTGATGGTGCA TGA CTGCGCAGATGAGAAAAAAGCCCTTCCTCCT	552
OY		481	CCAGAAAGATTTAAAAATTTCAAGTGTGGCCAAAAGACCTCTAGAGCCCCGCTTTAAGATTAT	540
Db		553	CCAGAAAGATTTAAAAATTTCAAGTGTGGCCAAAAGACCTCTAGAGCCCCGCTTTAAGATTAT	612
OY		541	GGGGGAGATTCACCAACCATCGAGAAACGAGCCCTGGTTGGCGGCATCTACAGAGAGGAC	600
Db		613	GGGGGAGATTCACCAACCATCGAGAAACGAGCCCTGGTTGGCGGCATCTACAGAGAGGAC	672
OY		601	CGGGGGGGGCTCTGTCACTACGTGTGTGAGGCAAGCTTCATCAGCCCTTGTCTGGGTGATC	660
Db		673	CGGGGGGGGCTCTGTCACTACGTGTGTGAGGCAAGCTTCATCAGCCCTTGTCTGGGTGATC	732
OY		661	AGCGGCACACACTGCTTCATTGATTACCCAAAAGAGAGGAGATCAATATGCTCACTGGGT	720
Db		733	AGCGGCACACACTGCTTCATTGATTACCC- AAGAAAGAGAGATCAATATGCTCACTGGGT	791
OY		721	CGCTCAAGGCTTAACCTCCAAACGCAAGGAGAGATGAATTTGAGGTGGAAAACTTAATC	780
Db		792	CGCTCAAGGCTTAACCTCCAAACGCAAGGAGAGATGAATTTGAGGTGGAAAACTTAATC	851
OY		781	CTACACAAAGGACTACAGC-GCTGACAGCGTTGCTACACACA	821
Db		852	CTACACAAAGGACTACAGCGGCGGTGAAGCGTTGCTTCACACA	893
<hr/>				
RESULT 7				
LOCUS	CF13J2052	826 bp	mRNA	linear EST 05-Jan-2003
DEFINITION	UT-HF-PQO-aws-d-23-0-UT.r1 NIH_MGC_215 Homo sapiens cDNA clone IMAGE:30553726 5', mRNA sequence.			
ACCESSION	CF13J2052			
VERSION	CF13J2052.1	GI:33214924		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 826)			
TITLE	Bonaldo,M.F., Lennon,G. and Soares,M.B.			
JOURNAL	Normalization and subtraction: two approaches to facilitate gene discovery			
MEDLINE	Genome Res. 6 (9), 791-806 (1996)			
PUBMED	97044477			
COMMENT	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bentso-soares@uiowa.edu Tissue Procurement: Mary Hendrix cDNA Library preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/humanmtl.html			

FEATURES	Seq primer: pyx-5.
SOURCE	Location/Qualifiers
	1. 826
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:30553726"
	/lab_source="Chondrosarcoma Lung Metastasis cell lines"
	/1.phosb="PH10B (T1 phage resistant)"
	/clone_1ib="NH_MGC_215"
	/note="Organ: Lung; Vector: pyx-Abs; Site.1: EcoR I; Site.2: Not I; The library was constructed according Bonaldo, Lemmon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pyx-Abs vector. The library tag sequence located between the Not I site and the polyA tail is GATTAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN	
Query Match	60.2%; Score 779.8; DB 14; Length 826;
Best Local Similarity	99.3%; Pred. No. 9.4e-202;
Matches 814; Conservative	0; Mismatches 3; Indels 3; Gaps 3
Py	225 TCACCTTTTACGAGGAAAGGCGACGACACCATGAGCGGCGCCTGCGTCCCTGGAA 284
Db	1 TCACCTTTTACGAGGAAAGGCGACACCATGAGCGGCGCCTGCGTCCCTGGAA 59
Py	285 CTCGCGCACTGCTCTTCAGCAAAAGTACATGCCACATCTGATCTTTCAGCTGGG 344
Db	60 CTCGCGCACTGCTCTTCAGCAAAAGTACATGCCACATCTGATCTTTCAGCTGGG 119
Py	345 CCTGGGGAACATTAATTCGACAGAAACCCAGAACCCGAGAGCCGCTGGTGTATGT 404
Db	120 CCTGGGGAACATTAATTCGACAGAAACCCAGAACCCGAGAGCCGCTGGTGTATGT 179
Py	405 GAGAGTGAGCCCTAAAGCGCGCTTGTCCAAAGTGCATGTGTGCATGACTGGCGAGATGAA 464
Db	180 GAGAGTGAGCCCTAAAGCGCGCTTGTCCAAAGTGCATGTGTGCATGACTGGCGAGATGAA 239
Py	465 AAAGCCCTCTCTCTCTCCAGAAATTAATTAATTCAGTGTGGCAAAAGACTCTGAGGCC 524
Db	240 AAAGCCCTCTCTCTCTCCAGAAATTAATTAATTCAGTGTGGCAAAAGACTCTGAGGCC 299
Py	525 CGGCTTTAAGTATATGGGGGAGAAATTCACACCAATTCGAGAACCGACCTGGTTGGCGG 584
Db	300 CGGCTTTAAGTATATGGGGGAGAAATTCACACCAATTCGAGAACCGACCTGGTTGGCGG 359
Py	585 CATCTACAGAGGACCGGGGGGAGCTGTCACTACGATGTGTGAGGACGCTCATGAG 644
Db	360 CATCTACAGAGGACCGGGGGGAGCTGTGTCACTACGATGTGTGAGGAGGCTCATGAG 419
Py	645 CCTTGTGGGTGATCAAGCGCACACATGCTTCATTGATTAACCAAGAGAGAGACTTA 704
Db	420 CCTTGTGGGTGATCAAGCGCACACATGCTTCATTGATTAACCAAGAGAGAGACTTA 479
Py	705 CATGCTTAACCTGGGTGCTCAAGGCTTAACTCCACACGCAAGGGGAGATGAAGTTTGA 764
Db	480 CATGCTTAACCTGGGTGCTCAAGGCTTAACTCCACACGCAAGGGGAGATGAAGTTTGA 539
Py	765 GGTGGAACCTAATCTTACACAGAGCTACAGCGCTGACAGCTTGTCTCCACCAAG 824
Db	540 GGTGGAACCTAATCTTACACAGAGCTACAGCGCTGACAGCTTGTCTCCACCAATGA 599
Py	825 CATGGCCTTGTGGAAGTCCGTTCCAAAGAGGGCAGGTGTGGGAGCCATCCCGGACTAT 884
Db	600 CATGGCCTTGTGGAAGTCCGTTCCAAAGAGGGCAGGTGTGGGAGCCATCCCGGACTAT 659
Py	885 ACAGACCATCTGCGCTGCTGATGTATTAACATCCCAAGTTTGGCAACAAGTGTGAGAT 944

Db	660	ACAGACCAATCTGCGCTCGCTCGATCTATTAACAGATCCCGAGTTTGACACAAGCTGTGAGAT	719		
Qy	945	CACGTGCGCTTTGGAAAAAGAAATTCTACCGCATCTCTATCTCGGAGCAGCTGAAATATAC	1004		
Db	720	CACGTGCGCTTTGGAAAAAGAG-ATTCTACCGCATCTCTATCTCGGAGCAGCTGAAATATAC	778		
Qy	1005	TGTTGTGAAGCTGATTTCCCAACCGGAGAGTGTGAGAGGCC	1044		
Db	779	TGTTGTGAAGCTGATTTCCCA-CGGAGAGTGTGAGAGGCC	817		
RESULT 8	CF132539	827 bp	mRNA	linear	EST 05-AUG-2003
LOCUS	CF132539	UI-HF-Q0-awo-f-04-0-UI.r1 NIH_MGC_215	Homo sapiens	CDNA clone	
DEFINITION	IMAGE:30559899 5', mRNA sequence.				
ACCESSION	CF132539				
VERSION	CF132539.1	GI:33215909			
KEYWORDS	EST				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 827)				
	Bonaldi, M.F., Lennon, G. and Soares, M.B.				
	Normalization and subtraction: two approaches to facilitate gene				
	discovery				
JOURNAL	Genome Res.	6 (9),	791-806	(1996)	
MEDLINE	97044477				
PUBMED	8889548				
COMMENT	Contact: Soares, MB				
	Coordinated Laboratory for Computational Genomics				
	University of Iowa				
	375 Newton Road,	4156	MEBRF, Iowa City, IA 52242, USA		
	Tel: 319 335 8250				
	Fax: 319 335 9565				
	Email: bento-soares@iowa.edu				
	Tissue Procurement: Mary Hendrix				
	CDNA Library preparation: Dr. M. Bento Soares, University of Iowa				
	CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa				
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa				
	Clone Distribution: Distribution information can be found at				
	http://genome.uiowa.edu/distribution/humanfl.html				
	Seq primer: pYX-5.				
FEATURES	Location/Qualifiers				
source	1..827				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:30559899"				
	/tissue_type="Chondrosarcoma Lung Metastasis cell lines"				
	/lab_host="DH10B (TI phage resistant)"				
	/clone_id="NIH_MGC_215"				
	/note="Organ: Lung; Vector: pYX-Aac; Site 1: EcoR I;				
	Site 2: Not I; The library was constructed according				
	Bonaldi, Lennon and Soares, Genome Research, 6:791-806,				
	1996. Denatured RNA was size fractionated on a 1% agarose				
	gel. First strand cDNA synthesis was primed with oligo-dT				
	primer containing a Not I site. Double strand cDNA was				
	size selected according to mRNA size fraction, ligated				
	with EcoR I adaptor, digested with Not I and then cloned				
	directionally into pYX-Aac vector. The library tag				
	sequence located between the Not I site and the polyA tail				
	is GATAAGGCCA. Tissue was provided by Mary Hendrix."				
ORIGIN					
Query Match	60.0%;	Score 777.6;	DB 14;	Length 827;	
Best Local Similarity	98.3%;	Pred. No. 3.8e-201;			
Matches 816;	Conservative 0;	Mismatches 11;	Indels 3;	Gaps 3;	
Qy	225	TCACCTTACGAGAAAGCGCAGACATGACCATGAGCGCGCCCTGCGCTGAGAA	284		
Db	1	TCACCTTACGAGAAAGG-CAGACATGACCATGAGCGCGCCCTGCGCTGAGAA	59		

QY	285	CTCTGCCACTGTCTCTTACGAAAGTACATGCCCAACAGATCTGATGCTCTTACAGCTGGG	344
Db	60	CTCTGCCACTGTCTCTTACGAAAGTACATGCCCAACAGATCTGATGCTCTTACAGCTGGG	119
QY	345	CCTGGGGAAACATTAATTACTTGACAGAAACCCAGAACCCGAGGCGACCTTGCTGCTATGT	404
Db	120	CCTGGGGAAACATTAATTACTTGACAGAAACCCAGAACCCGAGGCGACCTTGCTGCTATGT	179
QY	405	GCAGGTGGGCTTAAAGCCGCTTGTTCCAAAGATGCATGTGATATGATCTGCGAATGCAAA	464
Db	180	GCAGGTGGGCTTAAAGCCGCTTGTTCCAAAGATGCATGTGATATGATCTGCGAATGCAAA	239
QY	465	AAAGCCCTCCTCCTCCACAAAGATTTAAATTTCAAGTGTGGCCAAAAGACTCTGAGGCGC	524
Db	240	AAAGCCCTCCTCCTCCACAAAGATTTAAATTTCAAGTGTGGCCAAAAGACTCTGAGGCGC	299
QY	525	CCGCTTTAAGATTAATTTGGGGGAGAAATTCACCAACATCGAGAACACGCCCTTGTCGGC	584
Db	300	CCGCTTTAAGATTAATTTGGGGGAGAAATTCACCAACATCGAGAACACGCCCTTGTCGGC	359
QY	585	CATCTACAGAGAGGACCGGGGGGGCTCTGTACCTTAAGTGTGTGAGAGGACCCCTCATCAG	644
Db	360	CATCTACAGAGAGGACCGGGGGGGCTCTGTACCTTAAGTGTGTGAGAGGACCCCTCATCAG	419
QY	645	CCCTTGTGGGTGATCAGACGCGCACACACTGCTTCATTGATTAACCCAAAGAGAGACTA	704
Db	420	CCCTTGTGGGTGATCAGACGCGCACACACTGCTTCATTGATTAACCCAAAGAGAGACTA	479
QY	705	CATCGTCTACCTGGGTCGCTCAAGGCTTTAACTCCAAACACGCAAGGGAGATGAATTTGA	764
Db	480	CATCGTCTACCTGGGTCGCTCAAGGCTTTAACTCCAAACACGCAAGGGAGATGAATTTGA	539
QY	765	GGTGGAAACCTAATTCCTTACACAGGACTACAGGCTGACACGCTTGCTACCAACACGA	824
Db	540	GGTGGAAACCTAATTCCTTACACAGGACTACAGGCTGACACGCTTGCTACCAACACGA	599
QY	825	CATTGCCCTTGTGAAGATCCGTTCCAAAGAGGAGGATGTGCGAGCATCCCGACTAT	884
Db	600	CATTGCCCTTGTGAAGATCCGTTCCAAAGAGGAGGATGTGCGAGCATCCCGACTAT	659
QY	885	ACAGACCATCTGCGCTCGCCCTGATGTATTAACGATCCCAAGTTTGGCACAGCTGTGAT	944
Db	660	ACAGACCATCTGCGCTCGCCCTGAT-TANTACGATCCCAAGTTTGGCACAGCTGTGAT	718
QY	945	CACGCGCTTGTGAAAGAAAGAAATTCCTACCGACTATCTATCCGAGAGACGTGAAATGAC	1004
Db	719	CACGCGCTTGTGAAAGAAAG-ATTCTACCGACTATCTATCCGAGAGACGTGAAATGAC	777
QY	1005	TGTTGTGAAGCTGATTTCCCAACCGGAGGTGCACAGCCCACTACTACG	1054
Db	778	TGTTGTGAAGCTGATTTCCCAACCGGAGGTGCACAGCCCACTACTACG	827
RESULT 9			
CF131558		814 bp	linear
LOCUS		UI-HF-F00-awg-P-06-0-UI.r1.NIH.MGC_215	EST 05-AUG-2003
DEFINITION		IMAGE:30558989 5', mRNA sequence.	
ACCESSION		CF131558	
VERSION		CF131558.1	GI:33213934
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE		1 (bases 1 to 814)	
		Bonaldo,M.F., Lennon,G. and Soares,M.B.	
		Normalization and subtraction: two approaches to facilitate gene	
		discovery	
JOURNAL		Genome Res. 6 (9), 791-806	(1996)
MEDLINE		97044477	
PubMed		8889548	

COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu

Tissue Procurement: Mary Hendrix
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/humanfl.html>
Seq primer: pYX-5.

FEATURES

Source

Location/Qualifiers

1..814
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30558989"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="NIH_MGC_215"
/note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 59.9%; Score 776.8; DB 14; Length 814;
Best Local Similarity 99.3%; Pred. No. 6.2e-201;
Matches 811; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

224 GTCACTTTTACCGGAAAGGCCGACATGACATGAGGCGGCGCTGCTGCTGCA 283
|||||
1 GTCACTTTTACCGGAAAGGCCGACATGACATGAGGCGGCGCTGCTGCTGCA 59

284 ACTTGCCACTGTCCTTCAGCAAACTACATGCGCCACAGATGATGCTTCAGCTG 343
60 ACTTGCCACTGTCCTTCAGCAAACTACATGCGCCACAGATGATGCTTCAGCTG 119

344 GCGTGGGAAACATTAATTAATGACGAAACCGGAGCGGACCGCTGCTGCTATG 403
|||||
120 GCGTGGGAAACATTAATTAATGACGAAACCGGAGCGGACCGCTGCTGCTATG 179

404 TGCAGTGGGCGCTTAAAGCCGCTGTCAGAGATGATGATGATGCTGGGAGTGA 463
180 TGCAGTGGGCGCTTAAAGCCGCTGTCAGAGATGATGATGATGCTGGGAGTGA 239

464 AAAAGCCCTCTCTCTCCAGAAAGATTAATTAATGATGCGCAAAAGACTGAGGC 523
240 AAAAGCCCTCTCTCTCCAGAAAGATTAATTAATGATGCGCAAAAGACTGAGGC 299

524 CCGGCTTTAAGATTATGAGGAGGAAATTCACCAATCGAAGACCGCTGTTGGCG 583
300 CCGGCTTTAAGATTATGAGGAGGAAATTCACCAATCGAAGACCGCTGTTGGCG 359

584 CCACTACAGAGAGCAACCGGAGGAGGCTGTGTCACTACGATGATGAGGAGGCTCATCA 643
360 CCACTACAGAGAGCAACCGGAGGAGGCTGTGTCACTACGATGATGAGGAGGCTCATCA 419

644 GCGCTTGTGTGTGATGAGGCAACACTGCTTCATTTATTAATCCCAAGAGAGAGACT 703
420 GCGCTTGTGTGTGATGAGGCAACACTGCTTCATTTATTAATCCCAAGAGAGAGACT 479

704 ACATCGCTACCTGGGTGCTCAAGCTTAACCTCAACACGCAAGGAGGAGTGAAGTTTG 763

Db

480 ACATCGCTACCTGGGTGCTCAAGCTTAACCTCAACACGCAAGGAGGAGTGAAGTTTG 539

Qy

764 AGGTGAAAACCTTAATCTTACACAGAGATTACAGGCTGACACGCTGCTACCAACAG 823

Db

540 AGGTGAAAACCTTAATCTTACACAGAGATTACAGGCTGACACGCTGCTACCAACAG 599

Qy

824 ACATGCTTGTGCTGAAGATCCGTTCCAGAGAGGAGGATGTCGAGCATCCGAGACTA 883

Db

600 ACATGCTTGTGCTGAAGATCCGTTCCAGAGAGGAGGATGTCGAGCATCCGAGACTA 659

Qy

884 TACAGACCATCTGCTGCTGCTGATGATTAACGATCCCAAGTTTGGCAACAGCTGAGA 943

Db

660 TACAGACCATCTGCTGCTGCTGATGATTAACGATCCCAAGTTTGGCAACAGCTGAGA 719

Qy

944 TCACCTGCTTTGAAAAGAAATTCACGACTATCTATCCGAGACGCTGAAAATGA 1003

Db

720 TCACCTGCTTTGAAAAGAAATTCACGACTATCTATCCGAGACGCTGAAAATGA 778

Qy

1004 CTGTTGTGAAGCTGATTTCCACCGGAGTGTACGA 1040

Db

779 CTGTTGTGAAGCTGATTTCCACCGGAGTGTACGA 814

RESULT 10
BM920114
LOCUS
DEFINITION
BM920114 1067 bp mRNA linear EST 12-MAR-2002
AGENCOURT 6706522 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5749987
5', mRNA sequence.
ACCESSION
BM920114
VERSION
BM920114.1 GI:19370493
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1067)
NIH-MGC <http://mgi.nci.nih.gov/>.
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
[found through the I.M.A.G.E. Consortium/LNL at:](http://image.llnl.gov)
<http://image.llnl.gov>
Plate: L1AM12780 row: d column: 20
High quality sequence stop: 754.
Location/Qualifiers
1..1067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5749987"
/lab_host="DH10B"
/clone_id="NIH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector: pCW-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 59.6%; Score 772.4; DB 12; Length 1067;
Best Local Similarity 93.5%; Pred. No. 1.2e-199;
Matches 880; Conservative 0; Mismatches 48; Indels 13; Gaps 7;

```

QY      1 ATGAGAGCCCTGTGGGCGGCTTCTCTGTGCTCGTGAAGCACTCCAAAGC 60
DB      127 ATGAGAGCCCTGTGGGCGGCTTCTCTGTGCTCGTGAAGCACTCCAAAGC 186
QY      61 AGCAATGAATTCATCAAGTTCATGCACTGTGACTGTCTAATGAGAGACATGTGTG 120
DB      187 AGCAATGAATTCATCAAGTTCATGCACTGTGACTGTCTAATGAGAGACATGTGTG 246
QY      121 TCCAAACAAGTACTTCTCCAACTTCACTGTGTCACTGCCCAAGAAATTCGAGGGCAG 180
DB      247 TCCAAACAAGTACTTCTCCAACTTCACTGTGTCACTGCCCAAGAAATTCGAGGGCAG 306
QY      181 CACTGTGAATATAGTAAAGTAAACCTGTATGAGGGGAAATGTCACTTTTACGAGGA 240
DB      307 CACTGTGAATATAGTAAAGTAAACCTGTATGAGGGGAAATGTCACTTTTACGAGGA 366
QY      241 AAGGCCAGCACTGACACATGAGGCGGCGCTGCTGCTGCAACTCTGCACTGTCTT 300
DB      367 AAGGCCAGCACTGACACATGAGGCGGCGCTGCTGCTGCAACTCTGCACTGTCTT 426
QY      301 CAGCAAACTGATCCATGCCCAAGATCTGATGTCTTCACTGCTGCGGCAAACTAAT 360
DB      427 CAGCAAACTGATCCATGCCCAAGATCTGATGTCTTCACTGCTGCGGCAAACTAAT 486
QY      361 TACTGAGAGAACCCAGAACACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 420
DB      487 TACTGAGAGAACCCAGAACACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 546
QY      421 CGGCTGTCCAAAGAGTGCATGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB      547 CGGCTGTCCAAAGAGTGCATGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
QY      481 CCAGAAAGATTTAAATTTGATGTGSCCAAAAGACTCTGAGGCGCGCTTTAAATTTAT 540
DB      607 CCAGAAAGATTTAAATTTGATGTGSCCAAAAGACTCTGAGGCGCGCTTTAAATTTAT 666
QY      541 GGGGGAGATTCACCAACCATTCAGAACACCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB      667 GGGGGAGATTCACCAACCATTCAGAACACCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 726
QY      601 CGGGGGGGGCTGTGTCACTACGTGTGTGAGGAGGAGCTTCATGAGGCTGTGAGTATC 660
DB      727 CGGGGGGGGCTGTGTCACTACGTGTGTGAGGAGGAGCTTCATGAGGCTGTGAGTATC 786
QY      661 AGCGGCACACACTGCTTCACTTATTTAACCAGAGAGAGACATTCGTCTA-CCTGGG 719
DB      787 AGCGGCACACACTGCTTCACTTATTTAACCAGAGAGAGACATTCGTCTA-CCTGGG 846
QY      720 TGGCTCAAGGCTTAATCTCCAAACGCGAAGG--GAGATGAAGTTGAGGTGAAAACT 776
DB      847 TGGCTCAAGGCTTAATCTCCAAACGCGCAGGGGAGATGAAGTTTGAAGTGGAAAACT 906
QY      777 AATCTCAACAAGAGCTACAGGCGC-TGACACGCTGTGTCAACAAGACATTTGAGCTT-- 833
DB      907 CATCTCAACAAGAGCTACAGGCGCTTGAACACGCTTGTCTACCAATGACATTTGAGCTT 966
QY      834 -GCTGAAGATCCGTTTCAAGAGAGGAGGAGTGT--GCGACAGCATCCCGAGCTAATAC 888
DB      967 CTGGAAGATCCGTTTCAAGAGAGGAGGAGTGTGTCANCAATCCCGAGATTAATACAA 1026
QY      889 ACCATTTGCTGCTGCTT-CGATGTATTAACGATCCCGAGTTTG 928
DB      1027 ANCATCTGCTGCTGCTCGAGAGATTAACGATCCCGCATTTG 1067

```

```

RESULT 11
LOCUS   BU501434      925 bp      mRNA      linear      EST 12-SEP-2002
DEFINITION
AGENCOURT 8877614 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6318152
5' mRNA sequence.
ACCESSION
BU501434.1 GI:22803985

```

```

KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1
AUTHORS   Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL    NIH-MGC http://imgc.nci.nih.gov/
COMMENT    National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-roma@nih.gov
           Tissue Procurement: ATCC
           CDNA Library Preparation: Rubin Laboratory
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LNL at:
           http://image.lnl.gov
           Plate: LCM2530 row: 0 column: 17
           High quality sequence stop: 663.
           Location/Qualifiers
             1. 925
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /mol_type="mRNA"
               /clone="IMAGE:6318152"
               /issue_type="normal pigmented retinal epithelium"
               /lab_host="PH10B (phage-resistant)"
               /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
               EcoRI; CDNA made by oligo-dT priming. Directionally
               cloned into EcoRI/XhoI sites using the following 5'
               adaptor: GGCACGAG(G). Library constructed by Ling Hong
               in the laboratory of Gerald M. Rubin (University of
               California, Berkeley) using ZAP-CDNA synthesis kit
               (Stratagene) and Superscript II RT (Life Technologies).
               Note: this is a NIH_MGC Library. |"

```

```

ORIGIN
Query Match      59.5%; Score 770.6; DB 13; Length 925;
Best Local Similarity 98.3%; Pred. No. 3.3e-199;
Matches 800; Conservative 0; Mismatches 10; Indels 4; Gaps 2;
QY      1 ATGAGAGCCCTGTGGGCGGCTTCTCTGTGCTCGTGAAGCACTCCAAAGC 60
DB      104 ATGAGAGCCCTGTGGGCGGCTTCTCTGTGCTCGTGAAGCACTCCAAAGC 163
QY      61 AGCAATGAATTCATCAAGTTCATGCACTGTGACTGTCTAATGAGAGACATGTGTG 120
DB      164 AGCAATGAATTCATCAAGTTCATGCACTGTGACTGTCTAATGAGAGACATGTGTG 223
QY      121 TCCAAACAAGTACTTCTCCAACTTCACTGTGTCACTGCCCAAGAAATTCGAGGGCAG 180
DB      224 TCCAAACAAGTACTTCTCCAACTTCACTGTGTCACTGCCCAAGAAATTCGAGGGCAG 283
QY      181 CACTGTGAATATAGTAAAGTAAACCTGTATGAGGGGAAATGTCACTTTTACGAGGA 240
DB      284 CACTGTGAATATAGTAAAGTAAACCTGTATGAGGGGAAATGTCACTTTTACGAGGA 343
QY      241 AAGGCCAGCACTGACACATGAGGCGGCGCTGCTGCTGCTGCAACTCTGCACTGTCTT 300
DB      344 AAGGCCAGCACTGACACATGAGGCGGCGCTGCTGCTGCAACTCTGCACTGTCTT 403
QY      301 CAGCAAACTGATCCATGCCCAAGATCTGATGTCTTCACTGCTGCGGCAAACTAAT 360
DB      404 CAGCAAACTGATCCATGCCCAAGATCTGATGTCTTCACTGCTGCGGCAAACTAAT 463
QY      361 TACTGAGAGAACCCAGAACACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 420
DB      464 TACTGAGAGAACCCAGAACACCGGAGGCGACCTGTGTCTATGTGACAGTGGGCTTAAG 523
QY      421 CGGCTGTCCAAAGAGTGCATGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

```



```
Db      541 CACCACCAAATGCTGTGTGCTGCTGACCCACAGTGGAAAAACAGATTCTGCCAGGGAGA 600
QY      1125 CTGAGGGGGAGACCCCTGCTGTCTGTTCCTCCAAAGCCCGATGACTTTGACTGGAAATTGTGAG 1184
Db      601 CTCAGGGGGAGACCCCTGCTGTGTTCCTCCAAAGCCCGATGACTTTGACTGGAAATTGTGAG 660
QY      1185 CTGGGGCCCGTGAATGTGCTGAAAGGACAAGCCAGGCGCTCTACACGAGAGTCTCACACTT 1244
Db      661 CTGGGGCCCGTGAATGTGCTGAAAGGACAAGCCAGGCGCTCTACACGAGAGTCTCACACTT 720
QY      1245 CTTAACCTTGATCCGAGTCACACCAAGGAAGAAATGGCCCTGCTGA 1296
Db      721 CTTAACCTTGATCCGAGTCACACCAAGGAAGAAATGGCCCTGCTGA 771
```

Search completed: March 21, 2004, 20:50:03
Job time : 3808 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:53:24 ; Search time 178.247 Seconds
(without alignments)
683.197 Million cell updates/sec

Title: US-10-076-421-2

Perfect score: 2394
Sequence: 1 MRRLARLLCVLVSDSKG.....VSHFLPWRSHKENGAL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp19808:*
- 2: geneseqp19908:*
- 3: geneseqp20008:*
- 4: geneseqp20018:*
- 5: geneseqp20028:*
- 6: geneseqp20038:*
- 7: geneseqp20038:*
- 8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2394	100.0	431	1 AAP50114	Aap50114 Sequence
2	2394	100.0	431	1 AAP60783	Aap60783 Human uro
3	2394	100.0	431	1 AAP70258	Aap70258 Sequence
4	2394	100.0	431	1 AAP80430	Aap80430 Deduced A
5	2394	100.0	431	1 AAP81204	Aap81204 Pro-uroki
6	2394	100.0	431	1 AAP92119	Aap92119 Natural h
7	2394	100.0	431	2 AAP07112	Aar07112 Human pro
8	2394	100.0	431	2 AAP04253	Aar04253 Human pro
9	2394	100.0	431	2 AAP63141	Aar63141 Pull leng
10	2394	100.0	431	2 AAP47903	Aar47903 Pro-uroki
11	2394	100.0	431	3 AAP50869	Aay50869 Human uro
12	2394	100.0	431	3 AAP95951	Aay95951 Human pla
13	2394	100.0	431	4 AAP84605	Aab84605 Amino aci
14	2394	100.0	431	5 AAP17128	Aae17128 Human uro
15	2394	100.0	431	5 AAP99228	Aau99228 Human pla
16	2394	100.0	431	5 AAP99460	Aag99460 sc-urp. 1
17	2394	100.0	431	6 AAP55855	Abt55855 Human uro
18	2394	100.0	431	6 ABUS5547	Abus5547 lung can
19	2394	100.0	431	6 ABUS6708	Abus6708 Lung can
20	2394	100.0	431	6 ABUI1076	Abui1076 Human uro
21	2394	100.0	431	6 ABR92137	Abr92137 Human cer
22	2394	100.0	434	2 AAP20537	Aar20537 Amidated
23	2394	100.0	434	2 AAP20538	Aar20538 Amidated
24	2394	100.0	436	2 AAR20536	Aar20536 Amidated
25	2392	99.9	431	7 ADE25745	Ade25745 Human pro

26	2391	99.9	431	1 AAP91886	Aap91886 Sequence
27	2391	99.9	431	1 AAP94764	Aap94764 Non-glyco
28	2391	99.9	431	6 AAP37128	Aae37128 Human uro
29	2391	99.9	431	7 ADD46429	Add46429 Human pro
30	2390	99.8	431	5 AAU99230	Aau99230 Human pla
31	2389	99.8	431	5 AAU99229	Aau99229 Human pla
32	2386	99.7	431	5 AAU99236	Aau99236 Human pla
33	2385	99.6	431	1 AAP30041	Aap30041 Sequence
34	2385	99.6	431	5 AAU99231	Aau99231 Human pla
35	2384	99.6	431	5 AAU99232	Aau99232 Human pla
36	2383	99.5	431	1 AAP70250	Aap70250 Sequence
37	2382	99.5	431	5 AAU99238	Aau99238 Human pla
38	2381	99.5	431	5 AAU99237	Aau99237 Human pla
39	2380	99.4	431	1 AAP60674	Aap60674 Modified
40	2380	99.4	431	5 AAU99234	Aau99234 Human pla
41	2379	99.4	431	5 AAU99233	Aau99233 Human pla
42	2378	99.3	431	2 AAU33198	Aay33198 Human uri
43	2377	99.3	431	2 AAU33199	Aay33199 Human uri
44	2377	99.3	431	5 AAU99239	Aau99239 Human pla
45	2376	99.2	431	2 AAU33121	Aay33121 Human uri

ALIGNMENTS

RESULT 1	
AAP50114	
ID AAP50114	standard; protein; 431 AA.
XX	
AC AAP50114;	
XX	
DT 27-SEP-1991	(first entry)
XX	
DE Sequence encoded by the signal sequence and noncoding region of the pro-	
DE UK structural gene (Sequence II).	
XX	
KW Enzyme; thrombosis therapy; embolic disease; single-chain pro-urokinase.	
XX	
OS Homo sapiens.	
XX	
PH Key	Location/Qualifiers
FT Peptide	1..20
FT	/label= signal peptide
FT Domain	21..177
FT	/label= A chain
FT Domain	179..431
FT	/label= B chain
XX	
PN EP154272-A.	
XX	
PD 11-SEP-1985.	
XX	
PF 23-FEB-1985;	85EP-00102031.
XX	
PR 27-FEB-1984;	84JP-00037119.
PR 31-JAN-1985;	85JP-00017969.
XX	
PA (GREC) GREEN CROSS CORP.	
XX	
PI Hiramatou R, Kaneda T, Nagai M, Rimura H, Nishida M, Suyama T;	
XX	
DR WPI; 1985-224693/37.	
DR N-PSDB; AAN50138.	
XX	
PT Glycosylated single-chain pro-urokinase - prepd. by cultivating animal	
PT cells transformed by DNA prepd. from m RNA.	
XX	
PS Disclosure; Page 8-10; 64pp; English.	
CC The inventors claim a method of producing single-chain pro-urokinase by	
CC using as template, mRNA obt. from cells of an established human kidney-	
CC derived cell line. The urokinase is used to treat thrombosis and embolic	
CC diseases as well as in the treatment of diseases in combination with	

CC anticancer agents
XX
SQ Sequence 431 AA;

Query Match 100.0%; Score 2394; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVLYVSDSKGSENLHQPNSNCDLNGTCVSNKYFNSNIHWNCNCPKKGQ 60
DB 1 MRALLARLLLCVLYVSDSKGSENLHQPNSNCDLNGTCVSNKYFNSNIHWNCNCPKKGQ 60
QY HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQGLGKN 120
DB HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQGLGKN 120
QY YCRPNDRRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEELKFCGQKTLRPRFKII 180
DB YCRPNDRRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEELKFCGQKTLRPRFKII 180
QY GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
DB GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
QY 181 GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
DB 181 GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
QY 241 RSRINSNTQGEKMEVENLIHKDYADTLAHNDIALKIKRSKGRCAQPSRTIQTICL 300
DB 241 RSRINSNTQGEKMEVENLIHKDYADTLAHNDIALKIKRSKGRCAQPSRTIQTICL 300
QY 301 PSMYNDPQFGTSCETIGFGENSTDYLYPEQLKMTVVKLISHRECOQPHYYSSEVTTKML 360
DB 301 PSMYNDPQFGTSCETIGFGENSTDYLYPEQLKMTVVKLISHRECOQPHYYSSEVTTKML 360
QY 361 CAADPQWKTDSCQDGGGGLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPMIR 420
DB 361 CAADPQWKTDSCQDGGGGLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPMIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

RESULT 2
AAP60783
ID AAP60783 standard; protein; 431 AA.
XX
AC AAP60783;
XX
DT 25-MAR-2003 (revised)
DT 23-OCT-1991 (first entry)
XX
DE Human urokinase.
XX
KW E.coli; high molecular urokinase.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 21..431
XX
XX JP61181377-A.
XX
XX
XX 14-AUG-1986.
XX
XX 25-JAN-1985; 85JP-00011032.
XX
XX 25-JAN-1985; 85JP-00011032.
XX
XX 25-JAN-1985; 85JP-00011032.
XX
XX (NISC) NISSAN CHEM IND LTD.
XX (HODO) HODOGAYA CHEM IND CO LTD.
XX (SAGA) SAGAMI CHEM RES CENTRE.
XX (CENG) CENTRAL GLASS CO LTD.
XX (NIPS) NIPPON SODA CO.
XX (TOYU) TOYO SODA MFG CO LTD.

XX
DR WPI: 1986-254744/39.
DR N-PSDB; AAP60703.
XX
XX
PT Human urokinase gene - has N-end of aminoacid sequence coded by codon
FT used in Escherichia coli.
XX
PS Disclosure; Fig 2; 19pp; Japanese.

CC The claimed gene product may be expressed in a transformed E.coli host,
CC for the efficient production of high molecular human urokinase. The N-
CC terminal of the protein expressed by the transforming plasmid is replaced
CC with a codon frequently used in E.coli. (Updated on 25-MAR-2003 to
CC correct PA field.)
XX

SQ Sequence 431 AA;
Query Match 100.0%; Score 2394; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVLYVSDSKGSENLHQPNSNCDLNGTCVSNKYFNSNIHWNCNCPKKGQ 60
DB 1 MRALLARLLLCVLYVSDSKGSENLHQPNSNCDLNGTCVSNKYFNSNIHWNCNCPKKGQ 60
QY HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQGLGKN 120
DB HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHRSDALQGLGKN 120
QY 121 YCRPNDRRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEELKFCGQKTLRPRFKII 180
DB 121 YCRPNDRRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEELKFCGQKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
DB 181 GGEFTTIENQPMFAIYRRHSGSVTVYCGSLISPCWVISATHCFTIDYPKKEDYIVYL 240
QY 241 RSRINSNTQGEKMEVENLIHKDYADTLAHNDIALKIKRSKGRCAQPSRTIQTICL 300
DB 241 RSRINSNTQGEKMEVENLIHKDYADTLAHNDIALKIKRSKGRCAQPSRTIQTICL 300
QY 301 PSMYNDPQFGTSCETIGFGENSTDYLYPEQLKMTVVKLISHRECOQPHYYSSEVTTKML 360
DB 301 PSMYNDPQFGTSCETIGFGENSTDYLYPEQLKMTVVKLISHRECOQPHYYSSEVTTKML 360
QY 361 CAADPQWKTDSCQDGGGGLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPMIR 420
DB 361 CAADPQWKTDSCQDGGGGLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPMIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

RESULT 3
AAP70258
ID AAP70258 standard; protein; 431 AA.
XX
XX AAP70258;
XX
AC AAP70258;
XX
XX 25-MAR-2003 (revised)
DT 19-MAY-1991 (first entry)
XX
XX
DE Sequence of human prourokinase and leader.
XX
KW Cardiovascular disease treatment; fibrin affinity; thrombolytic; enzyme;
XX protease.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT /label= leader

FT Protein 21.431
 FT /label= prourokinase
 XX EP231883-A.
 XX 12-AUG-1987.
 PD 29-JAN-1987; 87EP-00101209.
 XX 31-JAN-1986; 86JP-00017734.
 PR 30-JAN-1987; 87JP-00018626.
 XX (SAGA) SAGAMI CHEM RES CENTRE.
 PA (NIPS) NIPPON SODA CO.
 PA (CENG) CENTRAL GLASS CO LTD.
 PA (TOYO) TOYO SODA MFG CO LTD.
 PA (NISC) NISSAN CHEM IND LTD.
 PA (NISC) NISSAN CHEMICAL INDS KK.
 XX Tagawa M, Wada M, Yamada M, Yokoyama M, Numao N;
 PI WPI; 1987-222882/32.
 XX N-PSDB; AAN70390.
 DR Hybrid plasminogen activator-like polypeptide - having a region for
 PT affinity to fibrin from tissue plasminogen activator and a region from
 PT prourokinase.
 XX Disclosure; Fig 2(1-5); 64pp; English.
 PS The TPA portion of the claimed hybrid polypeptide (see FT) may consist of
 CC 2 kringles from N-terminal first serine to 219th glycine of human TPA, 1
 CC kringles from 188th serine to 219th glycine of human TPA or half a kringles
 CC from 161st methionine to 219th glycine (see AAP70257). The C-terminal
 CC half of the hybrid polypeptide may contain an AA sequence from 150th
 CC glutamine to C-terminal 411th leucine of prourokinase (see AAP70258).
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX Sequence 431 AA:
 SQ
 Query Match 100.0%; Score 2394; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3.1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALLRLLCVYVSDSGKSNELHVPNCDCCLNGSTCVSNKYFSFIHMCNPKKGGQ 60
 DB 1 MRALLRLLCVYVSDSGKSNELHVPNCDCCLNGSTCVSNKYFSFIHMCNPKKGGQ 60
 QY HCEIDSKTCYEGNGHFYRGKASTDTWGRPCLPMNSATVVOQTYHAHRSDALQLGKKN 120
 DB HCEIDSKTCYEGNGHFYRGKASTDTWGRPCLPMNSATVVOQTYHAHRSDALQLGKKN 120
 QY 121 YCNPDPNRRRPMCYVQGLKPLVQECVNDADGKSPSPPEELKFCQGGKTLRPRKII 180
 DB 121 YCNPDPNRRRPMCYVQGLKPLVQECVNDADGKSPSPPEELKFCQGGKTLRPRKII 180
 QY 121 YCNPDPNRRRPMCYVQGLKPLVQECVNDADGKSPSPPEELKFCQGGKTLRPRKII 180
 DB 121 YCNPDPNRRRPMCYVQGLKPLVQECVNDADGKSPSPPEELKFCQGGKTLRPRKII 180
 QY 181 GSEFTTIENPWFATYRRHRGSLVYVCGSLISPCWVLSATHCFIDYRKEDYIYLG 240
 DB 181 GSEFTTIENPWFATYRRHRGSLVYVCGSLISPCWVLSATHCFIDYRKEDYIYLG 240
 QY 241 RSLRLNTOGEMKEFEVENLILHSDYADTLAHNDIALALKIRSEKGCAPSRITQITCL 300
 DB 241 RSLRLNTOGEMKEFEVENLILHSDYADTLAHNDIALALKIRSEKGCAPSRITQITCL 300
 QY 301 PSMTYNDPQFSTCEITGFGKENSTDYLYPEQLKTTVYKLSHRBCQPHYGSEVTTKML 360
 DB 301 PSMTYNDPQFSTCEITGFGKENSTDYLYPEQLKTTVYKLSHRBCQPHYGSEVTTKML 360
 QY 361 CAADPQWKTDSCGDSGGLVCSLOGRMTLTGIYVWGRGALDKDPQVYRVSHPLPMIR 420
 DB 361 CAADPQWKTDSCGDSGGLVCSLOGRMTLTGIYVWGRGALDKDPQVYRVSHPLPMIR 420
 QY 421 SHTKEENGIAL 431

Db 421 SHTKEENGIAL 431
 RESULT 4
 AAP80430
 ID AAP80430 standard; protein; 431 AA.
 XX AAP80430;
 AC
 XX 25-MAR-2003 (revised)
 DT 14-SEP-1990 (first entry)
 XX Deduced AA sequence of the single chain urokinase plasminogen activator
 DE (SCU-PA) cDNA insert prepared from human Hep3 cells.
 XX Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells;
 KM glyceroldehyde-3-phosphate dehydrogenase gene; thrombosis prevention;
 XX thrombosis treatment.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Peptide 1..20
 FT Protein 21..411
 XX EP288435-A.
 PN 26-OCT-1988.
 PD 11-APR-1988; 88EP-00810234.
 XX 15-APR-1987; 87GB-00009081.
 PR 16-JUN-1987; 87GB-00014059.
 PR 04-DEC-1987; 87JE-00003299.
 XX (CIBA) CIBA GEIGY AG.
 XX
 PI Meyhack B, Helm J, Burgi R;
 XX WPI; 1988-301440/43.
 DR N-PSDB; AAN80981.
 XX
 PT Prod. of human single chain urokinase-type plasminogen activator - by
 PT culturing yeast strain transformed with hybrid vector cong. yeast
 PT expression control sequences.
 PS Example 1; Fig 2; 48pp; English.
 XX The patent is for the prodn. of human single chain urokinase-type
 CC plasminogen activator (uPA). Mutants of scu-PA are especially those
 CC which render the protein protease resistant. Such scu-PA mutants are
 CC covalently modified at sites of proteolysis by proteases occurring in
 CC blood such as thrombin or plasmin, so that they are no longer susceptible
 CC to protease hydrolysis at these locations. The target sites include
 CC Lys135 to Lys136 (cleavage at this site generates the so-called low
 CC molecular weight form of scu-PA or LUK), Arg156 to Phe157 (susceptible to
 CC thrombin attack) and Lys158 to Ile159 (cleavage at this site by plasmin
 CC generates tuc-PA). Suitable scu-PA mutants have site specific
 CC substitutions, insertions or deletions of residues at one or more of
 CC these target sites. Especially preferred are those mutants in which one
 CC amino acid residue or both amino acid residues forming the target sites
 CC are deleted or in which at least one of these amino acid residues is
 CC replaced by another amino acid residue so that the resulting mutants are
 CC resistant to proteolytic attack. The uPA proteins exhibit the biological
 CC activity of natural human uPA without any refolding procedure being
 CC necessary. They can be used as for known PAs in humans for the prevention
 CC or treatment of thrombosis or other conditions where it is desired to
 CC produce local fibrinolytic or proteolytic activity. (Updated on 25-MAR-
 CC 2003 to correct PR field.)
 XX Sequence 431 AA:
 SQ

Query Match 100.0%; Score 2394; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3.1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRALLARLLLCVTVVSDSGSNELHQPNSNCDLNGCTCVSNKYFSNTHMCNCPKKEGGQ 60
DB 1 MRALLARLLLCVTVVSDSGSNELHQPNSNCDLNGCTCVSNKYFSNTHMCNCPKKEGGQ 60
QY HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLTQTYHAHRSDALQGLGKRN 120
DB HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLTQTYHAHRSDALQGLGKRN 120
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLTQTYHAHRSDALQGLGKRN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLTQTYHAHRSDALQGLGKRN 120
QY 121 YCRNPNRRRPMWCYQVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
DB 121 YCRNPNRRRPMWCYQVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHGGSVTVYCGSLISPCWVIAATCFIDYPKKEDYIVYL 240
DB 181 GGEFTTIENQPMFAIYRRHGGSVTVYCGSLISPCWVIAATCFIDYPKKEDYIVYL 240
QY 241 RSRINSNTGEMKEFEVENTLHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICL 300
DB 241 RSRINSNTGEMKEFEVENTLHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICL 300
QY 301 PSMYNDPQFGTSCETITFGKENSVDLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML 360
DB 301 PSMYNDPQFGTSCETITFGKENSVDLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCQDGGSGPLVCSLQGRMTLTGIVSMGRGALKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPQWKTDSCQDGGSGPLVCSLQGRMTLTGIVSMGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

```

RESULT 5

AAP81204
 ID AAP81204 standard; protein; 431 AA.

AC AAP81204;

XX 25-MAR-2003 (revised)
 DT 03-DEC-1990 (first entry)

XX Pro-urokinase with signal sequence.

DE pro-urokinase (pro-UK); plasminogen activator; pUK3; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein /label= signal peptide

FT 21..431

XX EP265874-A.

XX 04-MAY-1988.

XX 23-OCT-1987; 87EP-00115600.

XX 23-OCT-1986; 86JP-00253078.

XX (GREC) GREEN CROSS CORP.

PI Anatsuji Y, Okabayashi K, Nagai M, Arimura H, Suyama T;

DR WPI, 1988-121000/18.

XX N-PSDB; AAN81558.

PT glycosylated single-chain pro-urokinase prodn. - by cultivating DHFR gene
 PT -deficient CHO-K1 cells transformed with a plasmid contg. CDNA, SV40
 promoter and DHFR gene.

XX Disclosure; Page 7; 19pp; English.

XX The Arg at position 2 is encoded by TGA (sic). Possible error in the
 CC specification. Should read CGA ? The pro-UK gene was derived from plasmid
 CC pUK3. The CDNA was synthesised using urokinase mRNA isolated from a
 CC human kidney cell line. Pro-UK was cloned into a SV40 promoter-contg.
 CC plasmid, down -stream of the promoter. This plasmid was then ligated to
 CC a DHFR- gene contg. plasmid so that pro-UK and DHFR are inserted in
 CC opposite directions. The recombinant plasmid was used to transform CHO-K1
 CC cell derived DHFR gene-deficient host cells to produce glycosylated
 CC single-chain pro-UK. (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 431 AA;

Query Match 100.0%; Score 2394; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3.1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRALLARLLLCVTVVSDSGSNELHQPNSNCDLNGCTCVSNKYFSNTHMCNCPKKEGGQ 60
DB 1 MRALLARLLLCVTVVSDSGSNELHQPNSNCDLNGCTCVSNKYFSNTHMCNCPKKEGGQ 60
QY HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLTQTYHAHRSDALQGLGKRN 120
DB HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLTQTYHAHRSDALQGLGKRN 120
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLTQTYHAHRSDALQGLGKRN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPMNSATVLTQTYHAHRSDALQGLGKRN 120
QY 121 YCRNPNRRRPMWCYQVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
DB 121 YCRNPNRRRPMWCYQVGLKPLVQECMVHDCADGKKSPPEELKFOCGQKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHGGSVTVYCGSLISPCWVIAATCFIDYPKKEDYIVYL 240
DB 181 GGEFTTIENQPMFAIYRRHGGSVTVYCGSLISPCWVIAATCFIDYPKKEDYIVYL 240
QY 241 RSRINSNTGEMKEFEVENTLHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICL 300
DB 241 RSRINSNTGEMKEFEVENTLHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICL 300
QY 301 PSMYNDPQFGTSCETITFGKENSVDLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML 360
DB 301 PSMYNDPQFGTSCETITFGKENSVDLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCQDGGSGPLVCSLQGRMTLTGIVSMGRGALKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPQWKTDSCQDGGSGPLVCSLQGRMTLTGIVSMGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

```

RESULT 6

AAP92119
 ID AAP92119 standard; protein; 431 AA.

XX AAP92119;

XX 25-MAR-2003 (revised)

DT 29-JUN-1990 (first entry)

XX Natural human prourokinase.

DE Human prourokinase; antithrombotic; derivative.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "Optional in new deriv."

```

FT Misc-difference 2.155
FT /note= "Incorporated into new deriv."
FT Misc-difference 135
FT /note= "May be replaced by a non-basic AA in new deriv."
FT Misc-difference 156
FT /note= "Undefined residue in new deriv."
FT Misc-difference 157
FT /note= "Pro, Gly, Ala or Val in new deriv."
FT Misc-difference 158
FT /note= "Lys or Arg in new deriv."
FT Misc-difference 158
FT /note= "Lys or Arg in new deriv."
PN MO8901513-A.
XX
XX
XX 23-FEB-1989.
XX
XX 18-AUG-1988; 88MO-JP000815.
XX
XX 19-AUG-1987; 87JP-00204149.
XX
XX (SAGA ) SAGAMI CHEM RES CENTRE.
XX (CENG ) CENTRAL GLASS CO LTD.
XX (HODO ) HODOGAYA CHEM KK.
XX (NIPS ) NIPPON SODA CO.
XX (NISC ) NISSAN CHEM IND LTD.
XX
XX Kobayashi Y, Omori M, Yamada C;
XX PI
XX WPI; 1989-068869/09.
XX DR N-PSDB; AAN91075.
XX
XX Antithrombotic fast-acting pro-urokinase deriv. - produced by culture of
XX PT E.coli transformant contg. new plasmid of PMUT9Q family.
XX PS
XX Disclosure; Fig 1; 75pp; Japanese.
XX
XX A human prourokinase (PU) deriv. is new which is based upon residues 2-
XX CC 155 of natural human prourokinase. The new deriv. is produced by E. coli
XX CC J103/PMUT9Q-RPK in culture. It is a fast-acting drug for the treatment
XX CC and prevention of thrombosis. (Updated on 25-MAR-2003 to correct PA
XX CC field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 431 AA;
SQ
Query Match 100.0%; Score 2394; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRLLARLLLCVIVSDSKGSEIHQVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFFGGQ 60
DB 1 MRLLARLLLCVIVSDSKGSEIHQVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFFGGQ 60
QY 61 HCEIDSKTCYBENGHFYRGKASTDTMGRPCLPWNSATYVLOQTYHAHRSALDGLGKHN 120
DB 61 HCEIDSKTCYBENGHFYRGKASTDTMGRPCLPWNSATYVLOQTYHAHRSALDGLGKHN 120
QY 121 YCNPNDNRPRPWCYVGVGLKPLVQECMVHDCADGKRPSPPEELKFOCGKTLRPRFKII 180
DB 121 YCNPNDNRPRPWCYVGVGLKPLVQECMVHDCADGKRPSPPEELKFOCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHGGSVTVVCGSLISPWCVISAHCFTIDYRKEDYIVYL 240
DB 181 GGEFTTIENQPMFAIYRRHGGSVTVVCGSLISPWCVISAHCFTIDYRKEDYIVYL 240
QY 241 RSLNLSNTQEMKFEVENLILHSDYADTLAHNDIALKIRSKREGCAOPSTIOTICL 300
DB 241 RSLNLSNTQEMKFEVENLILHSDYADTLAHNDIALKIRSKREGCAOPSTIOTICL 300
QY 301 PSYNDPQFSGTCEITGFGKENSTDYLYPQOLKMTVVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSYNDPQFSGTCEITGFGKENSTDYLYPQOLKMTVVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSMRGCCALDKQPGVTVRVSHFLPMIR 420
DB 361 CAADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSMRGCCALDKQPGVTVRVSHFLPMIR 420

```

```

DB 361 CAADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSMRGCCALDKQPGVTVRVSHFLPMIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431
RESULT 7
AAR07112
ID AAR07112 standard; protein; 431 AA.
XX
AC AAR07112;
XX
XX 25-MAR-2003 (revised)
DT 24-JAN-1991 (first entry)
XX
DE Human pro-Urokinase encoded by plasmid pUK1.
XX
XX pro-Urokinase; transgenic mice.
XX
XX Synthetic.
XX
XX EP390592-A.
XX
XX 03-OCT-1990.
XX
XX 30-MAR-1990; 90EP-00303445.
XX
XX 31-MAR-1989; 89DP-00078574.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX (EXPE-) CENT INST EXPER ANIMALS.
XX (JIKK-) JIKKEN DOBUTSU CHUO KENK.
XX
XX Sekine S, Ito S, Katsuki M;
XX PI
XX WPI; 1990-299492/40.
XX DR N-PSDB; AAO06049.
XX
XX Prodn. of recombinant protein, esp. human pro-urokinase - from milk of
XX PT transgenic animals using promoter of bovine alpha S1 casein chromosomal
XX PT gene.
XX
XX Example; Table 1; 55pp; English.
XX
XX E.coli strain C600SF8 was transformed with recombinant plasmid containing
XX CC de DNA derived from human pharynx cancer cell strain Detroit 562. 10000
XX CC colonies were screened and one positive clone was identified. Plasmid
XX CC pUK1 was isolated and found to contain the coding region and 3' non-
XX CC coding region of pro-UK downstream of Cys(41). Four silent substitutions
XX CC were identified c.f. Holmes et al., Biotechnology, vol.3, p.923 (1985) as
XX CC fellows: (sic) Asn(254), AAC to AAT; Leu(340), CTA to CTG; Pro(345), CCC
XX CC to CCA; Gln(346), CAA to CAG. See also AAO06045-006048 and AAO06392.
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 431 AA;
SQ
Query Match 100.0%; Score 2394; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRLLARLLLCVIVSDSKGSEIHQVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFFGGQ 60
DB 1 MRLLARLLLCVIVSDSKGSEIHQVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFFGGQ 60
QY 61 HCEIDSKTCYBENGHFYRGKASTDTMGRPCLPWNSATYVLOQTYHAHRSALDGLGKHN 120
DB 61 HCEIDSKTCYBENGHFYRGKASTDTMGRPCLPWNSATYVLOQTYHAHRSALDGLGKHN 120
QY 121 YCNPNDNRPRPWCYVGVGLKPLVQECMVHDCADGKRPSPPEELKFOCGKTLRPRFKII 180
DB 121 YCNPNDNRPRPWCYVGVGLKPLVQECMVHDCADGKRPSPPEELKFOCGKTLRPRFKII 180

```

QY 181 GGEFTTINQPMFAIYRRHSGSVTVYVCGSLISPCWVISAHCFTIDYPKKEDYIVYLG 240
 Db 181 GGEFTTINQPMFAIYRRHSGSVTVYVCGSLISPCWVISAHCFTIDYPKKEDYIVYLG 240
 QY 241 RSRINSNTOGEMKFEVENILHKDYASADTLAHNDIALKIRSKERCAQPSRTIQTICL 300
 Db 241 RSRINSNTOGEMKFEVENILHKDYASADTLAHNDIALKIRSKERCAQPSRTIQTICL 300
 QY 301 PSMYNDPQFGTSCETITGFGKENSSTDYLYPEOLKMTVVKLISHRECOQPHYGSSEVTTKML 360
 Db 301 PSMYNDPQFGTSCETITGFGKENSSTDYLYPEOLKMTVVKLISHRECOQPHYGSSEVTTKML 360
 QY 361 CAADPQWKTDSCQDGGSGPLVCSLOGRMTLTGIVSWRGCAALKDKPGVYTRVSHFLPWIR 420
 Db 361 CAADPQWKTDSCQDGGSGPLVCSLOGRMTLTGIVSWRGCAALKDKPGVYTRVSHFLPWIR 420
 QY 421 SHTKEENGIAL 431
 Db 421 SHTKEENGIAL 431

RESULT 8

AA04253
 ID AAR04253 standard; protein; 431 AA.
 AC AAR04253;

XX 25-MAR-2003 (revised)
 DT 12-SEP-1990 (first entry)

XX Human pro-urokinase from the cDNA of clone pCUK176.

XX Non-glycosylated; pro-urokinase; E. coli; Pcrp promoter; MS-2 RBS.

XX Synthetic.

XX EP365894-A.

XX 02-MAY-1990.

XX 06-OCT-1989; 89EP-00118586.

XX 11-OCT-1988; 88GB-00023833.

XX (FARM) FARMITALIA ERBA SPA CARLO.

XX Brandazza A, Sarmientos P, Orsini G;

XX WPI; 1990-133447/18.

XX DR N-PSDB; AAQ04107.

XX Non-glycosylated pro-urokinase prodn. - using E coli B strains and E coli

XX promoter Pcrp and Shine-Dalgarno sequence MS-2.

XX Disclosure; Page ?; -pp; English.

XX SRR residue at position 21 is the start of the mature PROUK. Non-
 glycosylated PROUK (MW 45KD) produced by E.coli B strain containing the
 sequence. . See also AAQ04101-07. (Updated on 25-MAR-2003 to correct PA
 field.)

XX Sequence 431 AA;

Query Match 100.0%; Score 2394; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3,1e-184;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVAVSDSKSGSNELHOVPSNCDLNGTCSNRYFSNINHCNCPKFGGQ 60
 Db 1 MRALLARLLLCVAVSDSKSGSNELHOVPSNCDLNGTCSNRYFSNINHCNCPKFGGQ 60
 QY 61 HCEIDSKTCEBNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHRSALQLGKXN 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 61 HCEIDSKTCEBNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHRSALQLGKXN 120
 QY 121 YCRNPNRRRRPMKYQVQGLPVOECVHDCAQDKKPSPEELKFOCGKTLRPRFKII 180
 Db 121 YCRNPNRRRRPMKYQVQGLPVOECVHDCAQDKKPSPEELKFOCGKTLRPRFKII 180
 QY 181 GGEFTTINQPMFAIYRRHSGSVTVYVCGSLISPCWVISAHCFTIDYPKKEDYIVYLG 240
 Db 181 GGEFTTINQPMFAIYRRHSGSVTVYVCGSLISPCWVISAHCFTIDYPKKEDYIVYLG 240
 QY 241 RSRINSNTOGEMKFEVENILHKDYASADTLAHNDIALKIRSKERCAQPSRTIQTICL 300
 Db 241 RSRINSNTOGEMKFEVENILHKDYASADTLAHNDIALKIRSKERCAQPSRTIQTICL 300
 QY 301 PSMYNDPQFGTSCETITGFGKENSSTDYLYPEOLKMTVVKLISHRECOQPHYGSSEVTTKML 360
 Db 301 PSMYNDPQFGTSCETITGFGKENSSTDYLYPEOLKMTVVKLISHRECOQPHYGSSEVTTKML 360
 QY 361 CAADPQWKTDSCQDGGSGPLVCSLOGRMTLTGIVSWRGCAALKDKPGVYTRVSHFLPWIR 420
 Db 361 CAADPQWKTDSCQDGGSGPLVCSLOGRMTLTGIVSWRGCAALKDKPGVYTRVSHFLPWIR 420
 QY 421 SHTKEENGIAL 431
 Db 421 SHTKEENGIAL 431

RESULT 9

AA063141
 ID AAR63141 standard; protein; 431 AA.
 AC AAR63141;

XX 25-MAR-2003 (revised)
 DT 09-JUN-1995 (first entry)

XX Full length human urokinase protein.

XX Human urokinase glycoproteins; cardiovascular diseases;

XX pulmonary embolism.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Sig peptide 1..20

XX Disulfide-bond 70..151

XX Disulfide-bond 91..133

XX Disulfide-bond 122..146

XX Disulfide-bond 168..299

XX Cleavage-site 179..180

XX /note="cleavage of this site produces a bioactive two
 chain form of urokinase"

XX Disulfide-bond 209..225

XX Disulfide-bond 217..288

XX Disulfide-bond 313..382

XX Disulfide-bond 345..361

XX Disulfide-bond 372..400

XX EP620279-A1.

XX 19-OCT-1994.

XX 14-APR-1983; 94EP-00104777.

XX 15-APR-1982; 82US-00368773.

XX 14-MAR-1983; 83US-00474930.

XX 14-APR-1983; 83EP-00103629.

XX (GETH) GENENTECH INC.

XX Heyneker HL, Holmes WE, Vehar GA;

XX WPI; 1994-318362/40.

DR N-PSDB; AA073483.

XX Prod. of human urokinase glycoproteins - using a recombinant expression
PT system used for the treatment of vascular diseases or conditions.

XX Claim 1; Fig 4; 41pp; English.

XX AA073483 is the cDNA sequence which encodes AAR63141 the full length
CC 5400 dalton human urokinase (UK) protein. This cDNA was used in the
CC construction of a plasmid capable of transforming either yeast or
CC vertebrate cells, enabling them to produce the 5400 dalton human UK
CC protein. The UK glycoprotein produced could then be used in the treatment
CC of cardiovascular diseases, including pulmonary embolism. The UK produced
CC using this method had the advantage of a specific activity towards fibrin
CC and extant thrombi, not demonstrated previously with UK isolated from
CC natural sources. (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct
CC PR field.)

XX Sequence 431 AA;

XX Query Match 100.0%; Score 2394; DB 2; Length 431;

XX Best Local Similarity 100.0%; Pred. No. 3.1e-184; Mismatches 0; Indels 0; Gaps 0;

XX Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVIVSDSKSGSNEHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFKGGQ 60

Db 1 MRALLARLLLCVIVSDSKSGSNEHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFKGGQ 60

QY 61 HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPMSATVLOQTYHAHRSALQLGAKHN 120

Db 61 HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPMSATVLOQTYHAHRSALQLGAKHN 120

QY 121 YCRNPDNRRRPMCVQVGLKPLVQECMWHDCADGKSSPPEELKFCGQKTLAPRFKII 180

Db 121 YCRNPDNRRRPMCVQVGLKPLVQECMWHDCADGKSSPPEELKFCGQKTLAPRFKII 180

QY 181 GGEFTTIENOPWPAIYRRHSGSVTVVCGGSLSPCVWISATHCFLDYRKEDYIYVLG 240

Db 181 GGEFTTIENOPWPAIYRRHSGSVTVVCGGSLSPCVWISATHCFLDYRKEDYIYVLG 240

QY 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAOPSRTIOTICL 300

Db 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAOPSRTIOTICL 300

QY 301 PSMYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYVGSVTTKML 360

Db 301 PSMYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYVGSVTTKML 360

QY 361 CAADPQKTDSCQDSSGGLVCSLOGRMTLTGIVSWRGCALDKDPGVYTRVSHFLPMIR 420

Db 361 CAADPQKTDSCQDSSGGLVCSLOGRMTLTGIVSWRGCALDKDPGVYTRVSHFLPMIR 420

QY 421 SHTKEENGALAL 431

Db 421 SHTKEENGALAL 431

RESULT 10

AAR47903 standard; protein; 431 AA.

AAR47903;

DT 13-JUL-1994 (first entry)

DE Pro-urokinase derivative.

KM Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic; factor.

XX Homo sapiens.

XX JF05336965-A.

XX 21-DEC-1993.

XX 17-OCT-1991; 91JP-00269615.

XX 17-OCT-1991; 91JP-00269615.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX WPI; 1994-030907/04.

XX N-PSDB; AA055772.

XX Novel human pro-urokinase derivs. having long half-life - with high
PT thrombolytic activity, useful for treatment of thrombosis.

PS Disclosure; Page 15-17; 29pp; Japanese.

XX Sequences (AA055771-72) are pro-urokinase derivatives. The products have
CC an inserted sugar moiety having an amino acid substituted, depleted or
CC inserted variant around the thrombin cleavage site. They also have a long
CC half-life allowing them to be used in the treatment of thrombosis

XX Sequence 431 AA;

XX Query Match 100.0%; Score 2394; DB 2; Length 431;

XX Best Local Similarity 100.0%; Pred. No. 3.1e-184; Mismatches 0; Indels 0; Gaps 0;

XX Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVIVSDSKSGSNEHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFKGGQ 60

Db 1 MRALLARLLLCVIVSDSKSGSNEHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFKGGQ 60

QY 61 HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPMSATVLOQTYHAHRSALQLGAKHN 120

Db 61 HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPMSATVLOQTYHAHRSALQLGAKHN 120

QY 121 YCRNPDNRRRPMCVQVGLKPLVQECMWHDCADGKSSPPEELKFCGQKTLAPRFKII 180

Db 121 YCRNPDNRRRPMCVQVGLKPLVQECMWHDCADGKSSPPEELKFCGQKTLAPRFKII 180

QY 181 GGEFTTIENOPWPAIYRRHSGSVTVVCGGSLSPCVWISATHCFLDYRKEDYIYVLG 240

Db 181 GGEFTTIENOPWPAIYRRHSGSVTVVCGGSLSPCVWISATHCFLDYRKEDYIYVLG 240

QY 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAOPSRTIOTICL 300

Db 241 RSRINSTOGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAOPSRTIOTICL 300

QY 301 PSMYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYVGSVTTKML 360

Db 301 PSMYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECOQPHYVGSVTTKML 360

QY 361 CAADPQKTDSCQDSSGGLVCSLOGRMTLTGIVSWRGCALDKDPGVYTRVSHFLPMIR 420

Db 361 CAADPQKTDSCQDSSGGLVCSLOGRMTLTGIVSWRGCALDKDPGVYTRVSHFLPMIR 420

QY 421 SHTKEENGALAL 431

Db 421 SHTKEENGALAL 431

RESULT 11

AAV50869 standard; protein; 431 AA.

AAV50869;

DT 24-FEB-2000 (first entry)

DE Human urokinase protein fragment.

KM Urokinase; human; thrombolytic agent; streptokinase; antigenic;
KW blood clot; heart attack; treatment.

```
XX OS Homo sapiens.
XX PN MO9957251-A2.
XX PD 11-NOV-1999.
XX PF 06-MAY-1999; 99WO-US010086.
XX PR 06-MAY-1998; 98US-0084392P.
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX PI Zhang XC, Lin X, Tang JUN,
XX DR WPI, 2000-052966/04.
XX PT New thrombolytic agents derived from modified humanized streptokinase,
XX PF useful for treating blood clot disorders.
XX PS Disclosure; Page 46-48; 55pp; English.
CC This invention describes a novel thrombolytic agent comprising
CC streptokinase where at least one nonessential portion has been modified.
CC The invention also describes a method of forming a thrombolytic agent
CC comprising determining a nonessential portion of streptokinase and
CC modifying the nonessential portion to render the resulting protein less
CC antigenic. The modified streptokinase is used to treat blood clot
CC disorders, such as heart attacks. The modified streptokinase has less
CC antigenicity than streptokinase but is still able to complex plasminogen
CC and lead to plasminogen activation. Modified streptokinase with the
CC nonessential portions removed or truncated simplify the molecule. Such
CC smaller proteins are cheaper and easier to produce. This sequence
CC represents a fragment of the human urokinase protein which is used in the
CC description of the method of the invention
XX SQ
Sequence 431 AA;
Query Match 100.0%; Score 2394; DB 3; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALLARLLLCVLYVSDSKSGNELHQVPSNCDCLNGGTCVSNKYPNSNIHMCNCPKFFGQ 60
DB 1 MRALLARLLLCVLYVSDSKSGNELHQVPSNCDCLNGGTCVSNKYPNSNIHMCNCPKFFGQ 60
QY HCEIDSKTCYEGNGHFYRGKASTDTMGPRCLPNNSATVLOQTYHAHRSDALQGLGKHN 120
DB HCEIDSKTCYEGNGHFYRGKASTDTMGPRCLPNNSATVLOQTYHAHRSDALQGLGKHN 120
QY YCRNPDNRRRPPWCYOVGLKPLVOECWVNDGADGKKSPPEELKFCGGOKTLRPRFKII 180
DB YCRNPDNRRRPPWCYOVGLKPLVOECWVNDGADGKKSPPEELKFCGGOKTLRPRFKII 180
QY YCRNPDNRRRPPWCYOVGLKPLVOECWVNDGADGKKSPPEELKFCGGOKTLRPRFKII 180
DB YCRNPDNRRRPPWCYOVGLKPLVOECWVNDGADGKKSPPEELKFCGGOKTLRPRFKII 180
QY GGEFTTLENQWFAIYRRHRGSSVTYVCCGSLSPCWVISAATCFIDYKKEVIYVYG 240
DB GGEFTTLENQWFAIYRRHRGSSVTYVCCGSLSPCWVISAATCFIDYKKEVIYVYG 240
QY GGEFTTLENQWFAIYRRHRGSSVTYVCCGSLSPCWVISAATCFIDYKKEVIYVYG 240
DB GGEFTTLENQWFAIYRRHRGSSVTYVCCGSLSPCWVISAATCFIDYKKEVIYVYG 240
QY RSRINSTOGEMKPEVENLILHKDYSADTLAHNDIALILKIRSEKCAOPSRITOTICL 300
DB RSRINSTOGEMKPEVENLILHKDYSADTLAHNDIALILKIRSEKCAOPSRITOTICL 300
QY RSRINSTOGEMKPEVENLILHKDYSADTLAHNDIALILKIRSEKCAOPSRITOTICL 300
DB RSRINSTOGEMKPEVENLILHKDYSADTLAHNDIALILKIRSEKCAOPSRITOTICL 300
QY PSMWINDQFGTSCITFTFGKENSVDYLYPEQLKMTVVKLISHRECQOPHYVGSSEVTTKML 360
DB PSMWINDQFGTSCITFTFGKENSVDYLYPEQLKMTVVKLISHRECQOPHYVGSSEVTTKML 360
QY CAADPQKTDSCQSDSGPLVCSLOGRMTLGTIVSWRGALXDKDPVTVRVSHPLPMIR 420
DB CAADPQKTDSCQSDSGPLVCSLOGRMTLGTIVSWRGALXDKDPVTVRVSHPLPMIR 420
QY SHTKEENGAL 431
DB SHTKEENGAL 431
QY SHTKEENGAL 431
DB SHTKEENGAL 431
```

```
RESULT 12
ID AAY99591 standard; protein; 431 AA.
XX AC AAY99591;
XX DT 13-SEP-2000 (first entry)
XX DE Human plasminogen activator urokinase, u-PA.
XX KW Human; serine protease; plasminogen activator; cardiac; thrombolytic;
XX heart attack; stroke; blood clotting disorder.
XX OS Homo sapiens.
XX PN W0200032759-A1.
XX PD 08-JUN-2000.
XX PF 06-MAY-1999; 99WO-US009991.
XX PR 02-DEC-1998; 98US-0110588P.
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX PI Lin X, Zhang XC, Tang JUN,
XX DR WPI, 2000-422975/36.
XX PT Polypeptide with plasminogen activator activity useful as thrombolytic
XX agent for treating blood clot disorders e.g. heart attack, comprises 10
XX amino acid peptide fragment for recognition or activation of plasminogen.
XX PS Disclosure; Page 26-28; 41pp; English.
CC The present sequence is human plasminogen activator urokinase (u-PA), a
CC serine protease which hydrolyses a peptide bond in human plasminogen
CC to convert it to its active form, plasmin. Plasminogen is the principal
CC serine protease zymogen in the extracellular fluids of vertebrates.
CC Plasmin is implicated in pericellular proteolysis associated with a wide
CC range of physiological and pathological processes. Plasminogen activators
CC regulate plasminogen expression either by hydrolysing a peptide bond, as
CC in the case of u-PA, or by forming tight binding complexes with
CC plasminogen to spontaneously convert it to plasmin. Review of sequence
CC homologues of several plasminogen activators and chymotrypsin has
CC identified a six amino acid peptide involved in plasminogen activation.
CC This peptide is particularly useful when inserted between amino acid
CC residues 644 and 645 of full length human plasminogen. Novel plasminogen
CC activators have been made based upon the plasminogen
CC activation/recognition site of plasminogen binding proteins. The
CC polypeptides are useful in preparing thrombolytic agents for treating
CC blood clotting disorders such as heart attack
XX SQ
Sequence 431 AA;
Query Match 100.0%; Score 2394; DB 3; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALLARLLLCVLYVSDSKSGNELHQVPSNCDCLNGGTCVSNKYPNSNIHMCNCPKFFGQ 60
DB 1 MRALLARLLLCVLYVSDSKSGNELHQVPSNCDCLNGGTCVSNKYPNSNIHMCNCPKFFGQ 60
QY HCEIDSKTCYEGNGHFYRGKASTDTMGPRCLPNNSATVLOQTYHAHRSDALQGLGKHN 120
DB HCEIDSKTCYEGNGHFYRGKASTDTMGPRCLPNNSATVLOQTYHAHRSDALQGLGKHN 120
QY HCEIDSKTCYEGNGHFYRGKASTDTMGPRCLPNNSATVLOQTYHAHRSDALQGLGKHN 120
DB HCEIDSKTCYEGNGHFYRGKASTDTMGPRCLPNNSATVLOQTYHAHRSDALQGLGKHN 120
QY YCRNPDNRRRPPWCYOVGLKPLVOECWVNDGADGKKSPPEELKFCGGOKTLRPRFKII 180
DB YCRNPDNRRRPPWCYOVGLKPLVOECWVNDGADGKKSPPEELKFCGGOKTLRPRFKII 180
```

QY 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKEDYIYVLG 240
 DB 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKEDYIYVLG 240
 QY 241 RSRINSTQGEEMKFEVENLILHKDYSADTLAHNDIALILKIRSEKCAQPSRTIQTICL 300
 DB 241 RSRINSTQGEEMKFEVENLILHKDYSADTLAHNDIALILKIRSEKCAQPSRTIQTICL 300
 QY 301 PSMTNDPQFSGTCEITFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
 DB 301 PSMTNDPQFSGTCEITFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
 QY 361 CAADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSWGRCALDKDKPGVYTRVSHFLPMIR 420
 DB 361 CAADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSWGRCALDKDKPGVYTRVSHFLPMIR 420
 QY 421 SHTKEENGIAL 431
 DB 421 SHTKEENGIAL 431

RESULT 13

AAB84605
 ID AAB84605 standard; protein; 431 AA.

AC AAB84605;

DT 05-SEP-2001 (first entry)

DE Amino acid sequence of urokinase plasminogen activator.

XX Growth factor; protein inhibitor; protease; damaged tissue;
 KM platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
 KM connective tissue derived growth factor; CTGF; chrysalin; VEGF;
 KM keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
 KM transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
 KM granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
 KM vascular endothelial growth factor; urokinase plasminogen activator;
 KM dermal ulcer; wound.

XX Homo sapiens.

PN MO200149309-A2.

PD 12-JUL-2001.

PF 21-DEC-2000; 2000MO-IB001935.

PR 29-DEC-1999; 99GB-00030768.

XX (PF12) PFIZER LTD.

PA (PF12) PFIZER INC.

PI Davies MJ, Huggins JP, McIntosh FS, Occlleston NL;

DR WPI; 2001-418351/44.

DR N-PSDB; AAH28220.

PT Composition for the treatment of damaged tissue i.e. chronic wounds and
 PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
 PT factor.

PS Disclosure; Page 550; 572pp; English.

XX The specification describes a pharmaceutical composition, comprising a
 CC growth factor, an inhibitor agent, i.e. a protease. The inhibitor agent
 CC inhibits the action of at least one specific adverse protein, i.e. a
 CC protease, that is upregulated in a damaged tissue such as a wound
 CC environment. Growth factors which are included in the composition of the
 CC invention are platelet-derived growth factor (PDGF), fibroblast growth
 CC factor (FGF), connective tissue derived growth factor (CTGF),
 CC keratinocyte-derived growth factor (KGF), transforming growth factor-beta
 CC (TGF-beta), granulocyte macrophage colony stimulating factor (GM-CSF),

CC epidermal growth factor (EGF), vascular endothelial growth factor (VEGF),
 CC and chrysalin. Inhibitors which are included in the composition of the
 CC invention include inhibitors of urokinase-type plasminogen activator
 CC (uPA) and matrix metalloproteinase (MMP). The composition is useful for
 CC the treatment of chronic damaged tissue, i.e. wounds and dermal ulcers.
 CC The present sequence represents a human uPA, and is used to produce the
 CC composition of the invention

XX Sequence 431 AA;

Query Match 100.0%; Score 2394; DB 4; Length 431;

Best Local Similarity 100.0%; Pred. No. 3,1e-184; Indels 0; Gaps 0;

Matches 431; Conservative 0; Mismatches 0;

QY 1 MRALLARLLLCVIVSPDSKSNELHQPNSDCDCLNGSTCVSNKYFSNIHMCNCPKKGQ 60
 DB 1 MRALLARLLLCVIVSPDSKSNELHQPNSDCDCLNGSTCVSNKYFSNIHMCNCPKKGQ 60
 QY 61 HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNSATVLOQTYHAHRSALQGLAKEN 120
 DB 61 HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNSATVLOQTYHAHRSALQGLAKEN 120
 QY 121 YCRNPDRRRPNCYVQGLKPLVQECMVHDCAGKSSPPEELKPCGQKTLRPRKII 180
 DB 121 YCRNPDRRRPNCYVQGLKPLVQECMVHDCAGKSSPPEELKPCGQKTLRPRKII 180
 QY 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKEDYIYVLG 240
 DB 181 GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKEDYIYVLG 240
 QY 241 RSRINSTQGEEMKFEVENLILHKDYSADTLAHNDIALILKIRSEKCAQPSRTIQTICL 300
 DB 241 RSRINSTQGEEMKFEVENLILHKDYSADTLAHNDIALILKIRSEKCAQPSRTIQTICL 300
 QY 301 PSMTNDPQFSGTCEITFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
 DB 301 PSMTNDPQFSGTCEITFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
 QY 361 CAADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSWGRCALDKDKPGVYTRVSHFLPMIR 420
 DB 361 CAADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSWGRCALDKDKPGVYTRVSHFLPMIR 420
 QY 421 SHTKEENGIAL 431
 DB 421 SHTKEENGIAL 431

RESULT 14

AAB17128
 ID AAB17128 standard; protein; 431 AA.

AC AAB17128;

DT 18-APR-2002 (first entry)

DE Human uPA protein.

XX Human; cancer; urokinase-type plasminogen activator; uPA; inflammation;
 KM Ets-1 transcription factor; N-acetylglucosaminyltransferase V; Gnt-V;
 KM matrix-type metalloproteinase; MMP-1; MMP-3; gene therapy.

XX Homo sapiens.

PN MO200196606-A2.

PD 20-DEC-2001.

PF 14-JUN-2001; 2001MO-US019248.

PR 14-JUN-2000; 2000US-00593488.

XX (NYXI-) NYXIS NEURO THERAPIES INC.

PI Yamamoto H, Kroes R, Moskal JR;
 XX WPI; 2002-130746/17.
 DR N-PSDB; AAD27855.
 XX
 PT Identifying a compound for treating cancer, comprises detecting
 PT transcription factor Ets-1, N-acetylglucosaminyltransferase V, urokinase-
 PT type plasminogen activator, matrix-type metalloproteinase-1 and -3 gene
 PT expression.
 XX
 PS Example 1; Page 62-63; 63pp; English.
 XX
 CC The invention relates to a method of identifying a compound for treating
 CC cancer. The method involves detecting the expression of a panel of
 CC sequences selected from transcription factor Ets-1, urokinase-type
 CC plasminogen activator (uPA), N-acetylglucosaminyltransferase V (Gnt-V),
 CC matrix-type metalloproteinase (MMP)-1 and MMP-3 in the cell. The method
 CC is useful for identifying a compound that affects a cell, particularly a
 CC cancer cell or glioma cell, or a cell that is involved in inflammation.
 CC It is used for diagnosing and/or treating cancer or other conditions that
 CC are affected by one or more members of a panel of genes or their protein
 CC product. The method is also useful for drug discovery, drug safety
 CC evaluations and in gene therapy. The present sequence is human uPA
 CC protein
 CC
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 2394; DB 5; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3,1e-184; Mismatches 0; Gaps 0;
 Matches 431; Conservative 0; Indels 0; Gaps 0;
 QY 1 MRALLARLLLCVLVVS DSKSGSNE LHOVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGQ 60
 Db 1 MRALLARLLLCVLVVS DSKSGSNE LHOVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGQ 60
 QY 61 HCEIDSKTCTCYE GNGHFFYRGKASTDTMG R PCLP MNSATV LQO TYHAHRS D A LQ L G K R N 120
 Db 61 HCEIDSKTCTCYE GNGHFFYRGKASTDTMG R PCLP MNSATV LQO TYHAHRS D A LQ L G K R N 120
 QY 121 YCRNDRNR R R P M C V Y O V G L K P L V O E C M W H D C A D K K S S P P E E L K F O C G O K T L A R P R K I I 180
 Db 121 YCRNDRNR R R P M C V Y O V G L K P L V O E C M W H D C A D K K S S P P E E L K F O C G O K T L A R P R K I I 180
 QY 181 GGEFTT I E N O P W F A I Y R R H R G S V T Y V C G S L I S P C V I S A T H C F I D Y P K E D Y I Y L G 240
 Db 181 GGEFTT I E N O P W F A I Y R R H R G S V T Y V C G S L I S P C V I S A T H C F I D Y P K E D Y I Y L G 240
 QY 241 RSR L N S T O G E M K E V E N L I H K D Y S A D T L A H N D I A L M L K I R S K E G R C A O P S R T I Q T I C L 300
 Db 241 RSR L N S T O G E M K E V E N L I H K D Y S A D T L A H N D I A L M L K I R S K E G R C A O P S R T I Q T I C L 300
 QY 301 P S M N D P O F G T S C E I T G F G K E N S T D Y L Y P E O L K M T V K L I S H R C C O O P H Y Y G S E V T T K M L 360
 Db 301 P S M N D P O F G T S C E I T G F G K E N S T D Y L Y P E O L K M T V K L I S H R C C O O P H Y Y G S E V T T K M L 360
 QY 361 C A A D P O K M T D S C O G D S G G P L V C S I O G S M T L T G I V S W R G C A L K D K P G V Y T R V S H F L P W I R 420
 Db 361 C A A D P O K M T D S C O G D S G G P L V C S I O G S M T L T G I V S W R G C A L K D K P G V Y T R V S H F L P W I R 420
 QY 421 S H T K E E N G I A L 431
 Db 421 S H T K E E N G I A L 431
 RESULT 15
 ID AAU99228
 AAU99228 standard; protein; 431 AA.
 AC AAU99228;
 DT 24-SEP-2002 (first entry)
 XX Human plasminogen activator, urokinase (PLAU).

XX
 KM Human; Plasminogen activator; urokinase; PLAU; cancer; enzyme;
 KM cytostatic; serine protease; thrombolytic disorder; isogene;
 KM pulmonary embolism; chromosome 10q24-qter; haplotype; genotype; SNP;
 KM single nucleotide polymorphism; thrombolytic; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN W0200240503-A2.
 XX
 PD 23-MAY-2002.
 XX
 PF 14-NOV-2001; 2001WO-US044001.
 XX
 PR 17-NOV-2000; 2000US-0249703P.
 XX
 PA (GENA-) GENAISSANCE PHARM INC.
 PI Anastasio AE, Bentivegna SC, Koshy B;
 DR WPI; 2002-519370/55.
 DR N-PSDB; ABR6597, ABR6598.
 PT Genetic variants of Plasminogen activator, Urokinase (PLAU) isogenes,
 PT useful for improving efficiency and reliability in drug development for
 PT treating thrombolytic disorders and cancer.
 XX
 PS Claim 27; Fig 3; 92pp; English.
 XX
 CC The invention relates to a polynucleotide comprising a first nucleotide
 CC sequence (NS1) comprising a PLAU (plasminogen activator, urokinase, a
 CC serine protease) isogene selected from isogenes 1-9 and 11-20 given in
 CC the specification, where each isogene comprises the regions of the PLAU
 CC gene or cDNA and is further defined by the corresponding sequence of
 CC polymorphisms (defining single nucleotide polymorphisms, SNP). Also
 CC included are methods of haplotyping/genotyping (and predicting the
 CC haplotype/genotype of the PLAU gene of an individual, identifying an
 CC association between a trait and at least one haplotype or haplotype pair
 CC of the PLAU gene, an isolated oligonucleotide for detecting a
 CC polymorphism in the PLAU gene, a recombinant non-human organism
 CC transformed or transfected with the gene or cDNA, fragments of the
 CC polynucleotides of at least 10 base pairs encompassing a polymorphic
 CC site, an isolated polymorphic variant PLAU protein or fragment, an
 CC isolated monoclonal antibody specific for PLAU, a computer system for
 CC storing and analyzing polymorphism data for the PLAU gene and a genome
 CC anthology for the PLAU gene. PLAU is useful in screening for drugs
 CC targeting PLAU that are useful for treating thrombolytic disorders and
 CC cancers. The methods are useful for improving the efficiency and
 CC reliability of the discovery and development of drugs for treating
 CC diseases associated with PLAU activity, in validating PLAU as a drug
 CC target and in the design of clinical trials for treating a specific
 CC condition of disease associated with PLAU activity. The antibody is
 CC useful in diagnostic, prognostic and therapeutic methods. PLAU
 CC polynucleotides are useful in studying the expression and function of
 CC PLAU, and in expressing PLAU protein for use in screening for candidate
 CC drugs to treat diseases related to PLAU activity. The gene for PLAU is
 CC located on chromosome 10q24-qter. The present sequence represents the
 CC PLAU protein
 CC
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 2394; DB 5; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3,1e-184; Mismatches 0; Gaps 0;
 Matches 431; Conservative 0; Indels 0; Gaps 0;
 QY 1 MRALLARLLLCVLVVS DSKSGSNE LHOVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGQ 60
 Db 1 MRALLARLLLCVLVVS DSKSGSNE LHOVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGQ 60
 QY 61 HCEIDSKTCTCYE GNGHFFYRGKASTDTMG R PCLP MNSATV LQO TYHAHRS D A LQ L G K R N 120
 Db 61 HCEIDSKTCTCYE GNGHFFYRGKASTDTMG R PCLP MNSATV LQO TYHAHRS D A LQ L G K R N 120

Qy	121	YCRNPDNRRRPWCYVQVGLKPLVQECMVDHCDGSKKSSPPEELKFOCGQKTLRPRFKII	180
Db	121	YCRNPDNRRRPWCYVQVGLKPLVQECMVDHCDGSKKSSPPEELKFOCGQKTLRPRFKII	180
Qy	181	GGEFTTENQPMFAIYRHRGGSVTYVCGSLISPCWVISATHCFIDYPKEDYIYVLG	240
Db	181	GGEFTTENQPMFAIYRHRGGSVTYVCGSLISPCWVISATHCFIDYPKEDYIYVLG	240
Qy	241	RSRLNSTOGEEMKFEVENLIHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICL	300
Db	241	RSRLNSTOGEEMKFEVENLIHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICL	300
Qy	301	PSMNDPQFGTSCETITGFGKENS TDLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML	360
Db	301	PSMNDPQFGTSCETITGFGKENS TDLYPEQLKMTVVKLISHRECQOPHYGSEVTTKML	360
Qy	361	CAADPQWKTDSCQDSDGGLVCSIQGRMTLGTIVSMRGCALKDKPGVYTRVSHFLPMIR	420
Db	361	CAADPQWKTDSCQDSDGGLVCSIQGRMTLGTIVSMRGCALKDKPGVYTRVSHFLPMIR	420
Qy	421	SHTKEENGLAL 431	
Db	421	SHTKEENGLAL 431	

Search completed: March 18, 2004, 11:16:19
 Job time : 180.247 secs

THIS PAGE BLANK (USPTO)


```

QY 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLKLSHRECOQPHYYGSEVTTKML 360
DB 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLKLSHRECOQPHYYGSEVTTKML 360
QY 361 CAADPQWKTDSGQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPQWKTDSGQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTEKENGIAL 431
DB 421 SHTEKENGIAL 431

```

```

RESULT 2
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1986
; SEQ ID NO.1:
; LENGTH: 431
5188829-1

```

```

Query Match 100.0%; Score 2394; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 2,8e-198;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRALLARLLLCVLYVSDSKGSNELHQVPSNCDCLNGTGVSNKYFSNIHMCNCPKFGGQ 60
DB 1 MRALLARLLLCVLYVSDSKGSNELHQVPSNCDCLNGTGVSNKYFSNIHMCNCPKFGGQ 60
QY 61 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNASATVLOQTYHAHRSALQLGLGKN 120
DB 61 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNASATVLOQTYHAHRSALQLGLGKN 120
QY 121 YCRNPNRRRPMWCYVQGLKPLVQECMVDKADKSPPEELKQCGOKTLRPRFKII 180
DB 121 YCRNPNRRRPMWCYVQGLKPLVQECMVDKADKSPPEELKQCGOKTLRPRFKII 180
QY 181 GGEFTTIENQWPFPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVYL 240
DB 181 GGEFTTIENQWPFPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVYL 240
QY 241 RSRINNTQGEKMEFEVENLIHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICL 300
DB 241 RSRINNTQGEKMEFEVENLIHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLKLSHRECOQPHYYGSEVTTKML 360
DB 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLKLSHRECOQPHYYGSEVTTKML 360
QY 361 CAADPQWKTDSGQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPQWKTDSGQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTEKENGIAL 431
DB 421 SHTEKENGIAL 431

```

```

RESULT 3
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858

```

```

; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO.2:
; LENGTH: 430
5219569-2

```

```

Query Match 99.4%; Score 2378.5; DB 6; Length 430;
Best Local Similarity 99.8%; Pred. No. 6.1e-197;
Matches 430; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

QY 1 MRALLARLLLCVLYVSDSKGSNELHQVPSNCDCLNGTGVSNKYFSNIHMCNCPKFGGQ 60
DB 1 MRALLARLLLCVLYVSDSKGSNELHQVPSNCDCLNGTGVSNKYFSNIHMCNCPKFGGQ 60
QY 61 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNASATVLOQTYHAHRSALQLGLGKN 120
DB 61 HCEIDSKTCYEGNGHFRGKASTDTMGRPCLPWNASATVLOQTYHAHRSALQLGLGKN 120
QY 121 YCRNPNRRRPMWCYVQGLKPLVQECMVDKADKSPPEELKQCGOKTLRPRFKII 180
DB 121 YCRNPNRRRPMWCYVQGLKPLVQECMVDKADKSPPEELKQCGOKTLRPRFKII 180
QY 181 GGEFTTIENQWPFPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVYL 240
DB 181 GGEFTTIENQWPFPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVYL 240
QY 241 RSRINNTQGEKMEFEVENLIHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICL 300
DB 241 RSRINNTQGEKMEFEVENLIHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLKLSHRECOQPHYYGSEVTTKML 360
DB 301 PSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVLKLSHRECOQPHYYGSEVTTKML 360
QY 361 CAADPQWKTDSGQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPQWKTDSGQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTEKENGIAL 431
DB 421 SHTEKENGIAL 431

```

```

RESULT 4
US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:

```


NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: TS1108Cont.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)815-6508
TELEFAX: (404)815-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..19
OTHER INFORMATION: /label= peptide
OTHER INFORMATION: /note= "WAF signal"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 198..203
OTHER INFORMATION: /label= modified
OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

Query Match 96.6%; Score 2312.5; DB 1; Length 430;
Best Local Similarity 96.5%; Pred. No. 3e-191;
Matches 416; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

QY 1 MRALLARLLICVLVSDSKSNEHQVPSNCDCLNGTCVSNKYFSNIHWCNCPKFGGQ 60
DB 1 MRCILIS-LVIGLALAEVALASNEHQVPSNCDCLNGTCVSNKYFSNIHWCNCPKFGGQ 59
QY 61 HCEIDSKTCYEGNGHFRKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGAKHN 120
DB 60 HCEIDSKTCYEGNGHFRKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGAKHN 119
QY 121 YCNPNRRRPMCVQVGLKPLVQECVHDCADGKSSPPEELKFCGQKTLRPFKII 180
DB 120 YCNPNRRRPMCVQVGLKPLVQECVHDCADGKSSPPEELKFCGQKTLRPFKII 179
QY 181 GGEFTTINQWPAIYRRHGGSVTVVCGSLISPCWVISAHCFTIDPKEDYIVYLG 240
DB 180 GGEFTTINQWPAIYRRHGGSVTVVCGSLISPCWVISAHCFTIDPKEDYIVYLG 239
QY 241 RSLNSTQCEMKEVENLIHKDYADTLAHNDIALKIRSEKRCAPSRITQICL 300
DB 240 RSLNSTQCEMKEVENLIHKDYADTLAHNDIALKIRSEKRCAPSRITQICL 299
QY 301 PSMYNDPOFGTSCETIFGKENSTDYLYPQOLKMTVVKLISHRECOOPHYGSEVTKML 360
DB 300 PSMYNDPOFGTSCETIFGKENSTDYLYPQOLKMTVVKLISHRECOOPHYGSEVTKML 359
QY 361 CAADPOWKTDSGCGDSGGLVCSLQGRMTLTGIIVSWRGALCKDKPGVTVRVSHFLWIR 420
DB 360 CAADPOWKTDSGCGDSGGLVCSLQGRMTLTGIIVSWRGALCKDKPGVTVRVSHFLWIR 419
QY 421 SHTKEENGAL 431
DB 420 SHTKEENGAL 430

RESULT 5
US-08-087-163-1
Sequence 1, Application US/08087163
Patent No. 5472692
GENERAL INFORMATION:
APPLICANT: Liu, Jian-Ning
APPLICANT: Gurewicz, Victor
TITLE OF INVENTION: PRO-UKINASE MUTANTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,163
FILING DATE: 07/02/93
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04353/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
US-08-087-163-1

Query Match 96.1%; Score 2301; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-190;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SNEHQVPSNCDCLNGTCVSNKYFSNIHWCNCPKFGGQCHCEIDSKTCYEGNGHFRG 80
DB 1 SNEHQVPSNCDCLNGTCVSNKYFSNIHWCNCPKFGGQCHCEIDSKTCYEGNGHFRG 60
QY 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGAKHNVCNPNRRRPMCVQVGLK 140
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGAKHNVCNPNRRRPMCVQVGLK 120
QY 141 PIVOECVHDCADGKSSPPEELKFCGQKTLRPFKIIIGSEFTTINQWPAIYRRH 200
DB 121 PIVOECVHDCADGKSSPPEELKFCGQKTLRPFKIIIGSEFTTINQWPAIYRRH 180
QY 201 RGSSTVYVCGSLISPCWVISAHCFTIDPKEDYIVYLSRLNSTQCEMKEVENLI 260
DB 181 RGSSTVYVCGSLISPCWVISAHCFTIDPKEDYIVYLSRLNSTQCEMKEVENLI 240
QY 261 LHKDYADTLAHNDIALKIRSEKRCAPSRITQICLPSMYNDPOFGTSCETIFGK 320
DB 241 LHKDYADTLAHNDIALKIRSEKRCAPSRITQICLPSMYNDPOFGTSCETIFGK 300
QY 321 ENSTDYLYPQOLKMTVVKLISHRECOOPHYGSEVTKMLCAADPOWKTDSGCGSGL 380
DB 301 ENSTDYLYPQOLKMTVVKLISHRECOOPHYGSEVTKMLCAADPOWKTDSGCGSGL 360
QY 381 VCSLQGRMTLTGIIVSWRGALCKDKPGVTVRVSHFLWIRSHTEENGAL 431
DB 361 VCSLQGRMTLTGIIVSWRGALCKDKPGVTVRVSHFLWIRSHTEENGAL 411

RESULT 6
US-08-286-748B-18
Sequence 18, Application US/08286748B
Patent No. 575542
GENERAL INFORMATION:
APPLICANT: Victor Gurewicz
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY

```

; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Rasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-286-748B-18

Query Match          96.1%; Score 2301; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-190;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SNEHGVPSNDCCLNGGTCVSNKTFNSNHWNCNCKKFGGHCETDKSKTCYEGNGHFYRG 80
DB 1 SNEHGVPSNDCCLNGGTCVSNKTFNSNHWNCNCKKFGGHCETDKSKTCYEGNGHFYRG 60
QY 81 KASTDTMGPRCLPWNASATVLTQOTYHAHRSALQGLGHNHCRNDNRPRPWCYQVGLK 140
DB 61 KASTDTMGPRCLPWNASATVLTQOTYHAHRSALQGLGHNHCRNDNRPRPWCYQVGLK 120
QY 141 PLVDECMVHDCADGKSPPEELKFGCGKTLRPRFKIIGSEFTTIENQWPAIYRRH 200
DB 121 PLVDECMVHDCADGKSPPEELKFGCGKTLRPRFKIIGSEFTTIENQWPAIYRRH 180
QY 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPRKEDYIYLGSRINSNTQSEMKFEVENLI 260
DB 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPRKEDYIYLGSRINSNTQSEMKFEVENLI 240
QY 261 LHKDYSADTLAHNDIALKLRSEKRCAPSRITQITCLPSWYNDPQGTSCETIGFGK 320
DB 241 LHKDYSADTLAHNDIALKLRSEKRCAPSRITQITCLPSWYNDPQGTSCETIGFGK 300
QY 321 ENSTDYLYPBOCLKMTVYKLSHRECOQPHYGSEVTTKMLCAADPWKTDSCGDSGGPL 380
DB 301 ENSTDYLYPBOCLKMTVYKLSHRECOQPHYGSEVTTKMLCAADPWKTDSCGDSGGPL 360
QY 381 VCSLOGMVLTLGIYSWGRGALKDKRGVYTRVSHPLPWIRSHTEENGAL 431
DB 361 VCSLOGMVLTLGIYSWGRGALKDKRGVYTRVSHPLPWIRSHTEENGAL 411
```

RESULT 7
US-08-153-799-18

```

; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
; US-08-153-799-18

Query Match          96.1%; Score 2301; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-190;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SNEHGVPSNDCCLNGGTCVSNKTFNSNHWNCNCKKFGGHCETDKSKTCYEGNGHFYRG 80
DB 1 SNEHGVPSNDCCLNGGTCVSNKTFNSNHWNCNCKKFGGHCETDKSKTCYEGNGHFYRG 60
QY 81 KASTDTMGPRCLPWNASATVLTQOTYHAHRSALQGLGHNHCRNDNRPRPWCYQVGLK 140
DB 61 KASTDTMGPRCLPWNASATVLTQOTYHAHRSALQGLGHNHCRNDNRPRPWCYQVGLK 120
QY 141 PLVDECMVHDCADGKSPPEELKFGCGKTLRPRFKIIGSEFTTIENQWPAIYRRH 200
DB 121 PLVDECMVHDCADGKSPPEELKFGCGKTLRPRFKIIGSEFTTIENQWPAIYRRH 180
QY 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPRKEDYIYLGSRINSNTQSEMKFEVENLI 260
DB 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPRKEDYIYLGSRINSNTQSEMKFEVENLI 240
QY 261 LHKDYSADTLAHNDIALKLRSEKRCAPSRITQITCLPSWYNDPQGTSCETIGFGK 320
```

Db 241 LHKDYASDTLAHNDIALKIRSEKRCAPSRITQITCLPSMYNDPQFTSCETGFGK 300
 Qy 321 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSGCGDSGGL 380
 Db 301 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSGCGDSGGL 360
 Qy 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 431
 Db 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 411

RESULT 8
 US-09-403-736-2
 ; Sequence 2, Application US/09403736
 ; Patent No. 6638502
 ; GENERAL INFORMATION:

; APPLICANT: Aventis S.A.
 ; APPLICANT: LI, Hong
 ; APPLICANT: LI, He
 ; APPLICANT: GRISCELLI, Frank
 ; APPLICANT: OPOLO, Paule
 ; APPLICANT: SORIA, Claudine
 ; APPLICANT: RAGOT, Thierry
 ; APPLICANT: LERAND, Yves
 ; APPLICANT: SORIA, Jeanette
 ; APPLICANT: MABILLAT, Christelle
 ; APPLICANT: PERRICAUDET, Michel
 ; APPLICANT: YEH, Patrice
 ; TITLE OF INVENTION: Adenovirus-Mediated Intratumoral Delivery of An Angiogenesis Anta
 ; FILE REFERENCE: A2778A-US
 ; CURRENT APPLICATION NUMBER: US/09/403, 736
 ; PRIOR FILING DATE: 1999-10-26
 ; PRIOR APPLICATION NUMBER: PCT/EP98/02491
 ; PRIOR FILING DATE: 1998-11-05
 ; PRIOR APPLICATION NUMBER: 60/044, 980
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 411
 ; TYPE: PRT
 ; ORGANISM: humanuokinaae
 ; US-09-403-736-2

Query Match 96.0%; Score 2298; DB 4; Length 411;
 Best Local Similarity 99.8%; Pred. No. 5e-190;
 Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 21 SNEHQPNSNCDLNGGTCVSNKYFSNIHMCNCPKKGQHCIDSKTCYEGNGHFRYG 80
 Db 1 SNEHQPNSNCDLNGGTCVSNKYFSNIHMCNCPKKGQHCIDSKTCYEGNGHFRYG 60
 Qy 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKNHYCRNPNRRPWCYOVGLK 140
 Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKNHYCRNPNRRPWCYOVGLK 120
 Qy 141 PLVQECMVDHCDAGKRPSPPEELKFOCGOKTLRPRFKIIGSEFTTLENQPMFAIYRRH 200
 Db 121 PLVQECMVDHCDAGKRPSPPEELKFOCGOKTLRPRFKIIGSEFTTLENQPMFAIYRRH 180
 Qy 201 RGGSVTVVCGSLISPCWVISAATHCFIDYPKKEDYIYLGSRSLNSTOQEMKFEVENLI 260
 Db 181 RGGSVTVVCGSLISPCWVISAATHCFIDYPKKEDYIYLGSRSLNSTOQEMKFEVENLI 240
 Qy 261 LHKDYASDTLAHNDIALKIRSEKRCAPSRITQITCLPSMYNDPQFTSCETGFGK 320
 Db 241 LHKDYASDTLAHNDIALKIRSEKRCAPSRITQITCLPSMYNDPQFTSCETGFGK 300
 Qy 321 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSGCGDSGGL 380
 Db 301 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSGCGDSGGL 360

Qy 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 431
 Db 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 411

RESULT 9
 US-09-181-816-1
 ; Sequence 1, Application US/09181816
 ; Patent No. 6277818
 ; GENERAL INFORMATION:

; APPLICANT: MAZAR, Andrew P.
 ; APPLICANT: JONES, Terence R.
 ; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UNOKINASE
 ; FILE REFERENCE: 329042000300 SIDN 1-7
 ; CURRENT APPLICATION NUMBER: US/09/181, 816
 ; CURRENT FILING DATE: 1998-10-29
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 411
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-181-816-1

Query Match 95.7%; Score 2291; DB 3; Length 411;
 Best Local Similarity 99.8%; Pred. No. 2e-189;
 Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 SNEHQPNSNCDLNGGTCVSNKYFSNIHMCNCPKKGQHCIDSKTCYEGNGHFRYG 80
 Db 1 SNEHQPNSNCDLNGGTCVSNKYFSNIHMCNCPKKGQHCIDSKTCYEGNGHFRYG 60
 Qy 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKNHYCRNPNRRPWCYOVGLK 140
 Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKNHYCRNPNRRPWCYOVGLK 120
 Qy 141 PLVQECMVDHCDAGKRPSPPEELKFOCGOKTLRPRFKIIGSEFTTLENQPMFAIYRRH 200
 Db 121 PLVQECMVDHCDAGKRPSPPEELKFOCGOKTLRPRFKIIGSEFTTLENQPMFAIYRRH 180
 Qy 201 RGGSVTVVCGSLISPCWVISAATHCFIDYPKKEDYIYLGSRSLNSTOQEMKFEVENLI 260
 Db 181 RGGSVTVVCGSLISPCWVISAATHCFIDYPKKEDYIYLGSRSLNSTOQEMKFEVENLI 240
 Qy 261 LHKDYASDTLAHNDIALKIRSEKRCAPSRITQITCLPSMYNDPQFTSCETGFGK 320
 Db 241 LHKDYASDTLAHNDIALKIRSEKRCAPSRITQITCLPSMYNDPQFTSCETGFGK 300
 Qy 321 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSGCGDSGGL 380
 Db 301 ENSTDVLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSGCGDSGGL 360
 Qy 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 431
 Db 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEKENGAL 411

RESULT 10
 US-08-560-098A-48
 ; Sequence 48, Application US/08560098A
 ; Patent No. 5976841
 ; GENERAL INFORMATION:
 ; APPLICANT: WENDEL, Stephan
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; APPLICANT: STEFFENS, Gerd Josef
 ; TITLE OF INVENTION: Proteins having fibrinolytic and
 ; TITLE OF INVENTION: Coagulation-inhibiting Properties
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington

STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-48

Query Match 95.1%; Score 2277; DB 2; Length 411;
Best Local Similarity 99.3%; Pred. No. 3.2e-188;
Matches 408; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 21 SNEIHOVPSNDCCLNGGTCVSNKXFNIMHCNCKKFGGHCETDKSKTCYEGNGHYRG 80
DB 1 SNEIHOVPSNDCCLNGGTCVSNKXFNIMHCNCKKFGGHCETDKSKTCYEGNGHYRG 60
QY 81 KASTIDWGRPCLPMNSATVLOQTYHAHRSDALQGLGKKNYCRNPNRRRPMCVVQGLK 140
DB 61 KASTIDWGRPCLPMNSATVLOQTYHAHRSDALQGLGKKNYCRNPNRRRPMCVVQGLK 120
QY 141 PLVQECVHDCADCKKSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENOPWFAIYRRH 200
DB 121 PLVQECVHDCADCKKSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENOPWFAIYRRH 180
QY 201 RGGSVTVVCGGSLSPCWISATHCFIDYPKKEDIYVILGRSRINSTQGBMKFEVENLI 260
DB 181 RGGSVTVVCGGSLSPCWISATHCFIDYPKKEDIYVILGRSRINSTQGBMKFEVENLI 240
QY 261 LHKDYSADTLAHNDIALKIRSKGRCAOPSRITQITCLPSMYNDPQFGTSCGTTGFGK 320
DB 241 LHKDYSADTLAHNDIALKIRSKGRCAOPSRITQITCLPSMYNDPQFGTSCGTTGFGK 300
QY 321 ENSGDIYXPOLKMTVVKLIHRECOOPHYGSEVTTKMLCAADPQWKTSCGDSGGPL 380
DB 301 ENSGDIYXPOLKMTVVKLIHRECOOPHYGSEVTTKMLCAADPQWKTSCGDSGGPL 360
QY 381 VCSIQGEMTLTGIVSMRGCAKDKPGVYTRVSHFLPMIRSHTEENGILA 431
DB 361 VCSIQGEMTLTGIVSMRGCAKDKPGVYTRVSHFLPMIRSHTEENGILA 411

RESULT 11
US-08-560-098A-47
Sequence 47, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDEL, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties

NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKee, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-47

Query Match 84.7%; Score 2027.5; DB 2; Length 432;
Best Local Similarity 88.8%; Pred. No. 1.1e-166;
Matches 374; Conservative 11; Mismatches 19; Indels 17; Gaps 3;

QY 11 CVLVVDSKSGNELHVPNSCCCLNGGTCVSNKXFNIMHCNCKKFGGHCETDKSKTC 70
DB 29 CILVGSNGKN-----QCVTGEGRPKESHNGGFEETPEEY-----LQISKTC 71
QY 71 YEGNGHFYRGKASTDWMGRPCLPMNSATVLOQTYHAHRSDALQGLGKKNYCRNPNRRR 130
DB 72 YEGNGHFYRGKASTDWMGRPCLPMNSATVLOQTYHAHRSDALQGLGKKNYCRNPNRRR 131
QY 131 PMCVVQGLKPLVQECVHDCADCKKSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQ 190
DB 132 PMCVVQGLKPLVQECVHDCADCKKSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQ 191
QY 191 PMFAIYRRHRRGGSVTVVCGGSLSPCWISATHCFIDYPKKEDIYVILGRSRINSTQGB 250
DB 192 PMFAIYRRHRRGGSVTVVCGGSLSPCWISATHCFIDYPKKEDIYVILGRSRINSTQGB 251
QY 251 EMKFEVENLIHKKDYSADTLAHNDIALKIRSKGRCAOPSRITQITCLPSMYNDPQFG 310
DB 252 EMKFEVENLIHKKDYSADTLAHNDIALKIRSKGRCAOPSRITQITCLPSMYNDPQFG 311
QY 311 TSCGTTGFGKENSSTYLYPEOLKMTVVKLIHRECOOPHYGSEVTTKMLCAADPQWKT 370
DB 312 TSCGTTGFGKENSSTYLYPEOLKMTVVKLIHRECOOPHYGSEVTTKMLCAADPQWKT 371
QY 371 SCQDSGGPLVCSIQGEMTLTGIVSMRGCAKDKPGVYTRVSHFLPMIRSHTEENGILA 430
DB 372 SCQDSGGPLVCSIQGEMTLTGIVSMRGCAKDKPGVYTRVSHFLPMIRSHTEENGILA 431
QY 431 L 431
DB 432 L 432

RESULT 12

US-08-093-741-83
Sequence 83, Application US/08093741
Patent No. 5681721
GENERAL INFORMATION:
APPLICANT: STEFFENS, GERD J.
APPLICANT: WENDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITTING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
City: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,741
FILING DATE: 20-JUL-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ. ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-093-741-83

Query Match 84.5%; Score 2022; DB 1; Length 365;

Best Local Similarity 100.0%; Pred. No. 2.6e-166; Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKKNYCRNP 126
DB 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKKNYCRNP 60
QY 127 NRRRPMCYOVGLKPLVOECMVHDCADGKKSPPELKFQCGQKTLRPRFKIIGGEFTT 186
DB 61 NRRRPMCYOVGLKPLVOECMVHDCADGKKSPPELKFQCGQKTLRPRFKIIGGEFTT 120
QY 187 IENQPMFAIYRRHRGSGVTVYCGGSLISPCWISATHCFIDYPKKEDYIVYIGRSLNS 246
DB 121 IENQPMFAIYRRHRGSGVTVYCGGSLISPCWISATHCFIDYPKKEDYIVYIGRSLNS 180
QY 247 NTOGEMKFEVENILHKDYSADTLAHNDIALKIRSKGRCAOPRTTIOTICLPSPMYND 306
DB 181 NTOGEMKFEVENILHKDYSADTLAHNDIALKIRSKGRCAOPRTTIOTICLPSPMYND 240
QY 307 POGTSCETITGFGKSNSTDYLYPBOLMTVVKLISHRECOOPHYGSEVTTKMLCAADPQ 366
DB 241 POGTSCETITGFGKSNSTDYLYPBOLMTVVKLISHRECOOPHYGSEVTTKMLCAADPQ 300
QY 367 WKTDSCGSDGSGPLVCSLOGRMTLTGIVSGRGCAALKDKRGVYTRVSHPLPWRSHTKKE 426
DB 301 WKTDSCGSDGSGPLVCSLOGRMTLTGIVSGRGCAALKDKRGVYTRVSHPLPWRSHTKKE 360

QY 427 NGIAL 431
DB 361 NGIAL 365

RESULT 13
US-08-720-012-83
Sequence 83, Application US/08720012
Patent No. 5747291
GENERAL INFORMATION:
APPLICANT: STEFFENS, GERD J.
APPLICANT: WENDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITTING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
City: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ. ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-720-012-83

Query Match 84.5%; Score 2022; DB 1; Length 365;

Best Local Similarity 100.0%; Pred. No. 2.6e-166; Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKKNYCRNP 126
DB 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKKNYCRNP 60
QY 127 NRRRPMCYOVGLKPLVOECMVHDCADGKKSPPELKFQCGQKTLRPRFKIIGGEFTT 186
DB 61 NRRRPMCYOVGLKPLVOECMVHDCADGKKSPPELKFQCGQKTLRPRFKIIGGEFTT 120
QY 187 IENQPMFAIYRRHRGSGVTVYCGGSLISPCWISATHCFIDYPKKEDYIVYIGRSLNS 246
DB 121 IENQPMFAIYRRHRGSGVTVYCGGSLISPCWISATHCFIDYPKKEDYIVYIGRSLNS 180
QY 247 NTOGEMKFEVENILHKDYSADTLAHNDIALKIRSKGRCAOPRTTIOTICLPSPMYND 306
DB 181 NTOGEMKFEVENILHKDYSADTLAHNDIALKIRSKGRCAOPRTTIOTICLPSPMYND 240

QY 307 POFGTCSEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHYYSSEVTTKMLCAADPQ 366
| | | | |
Db 241 POFGTCSEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHYYSSEVTTKMLCAADPQ 300
QY 367 WKTSCQSGSGGPLVCSLQGRMTLTGIVSWGRGALKDKRQVYTRVSHFLPWIRSHKKEE 426
| | | | |
Db 301 WKTSCQSGSGGPLVCSLQGRMTLTGIVSWGRGALKDKRQVYTRVSHFLPWIRSHKKEE 360
QY 427 NGAL 431
| | | | |
Db 361 NGAL 365

RESULT 14

US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841

GENERAL INFORMATION:

APPLICANT: WENDEL, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/560,098A

FILING DATE: 17-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P 44 40 892.7

FILING DATE: 17-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 148/42448

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 393 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-560-098A-44

Query Match 84.5%; Score 2022; DB 2; Length 393;

Best Local Similarity 100.0%; Pred. No. 2.9e-166;

Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 SKTCYEGNGHFFRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQGLGKHNYCRND 126
| | | | |
Db 2 SKTCYEGNGHFFRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQGLGKHNYCRND 61
QY 127 NRRPWCYVQVGLKPLVQECMWHDCADGKRPSSPEELKFCQCGQKTLRPRFKIIGSEFTT 186
| | | | |
Db 62 NRRPWCYVQVGLKPLVQECMWHDCADGKRPSSPEELKFCQCGQKTLRPRFKIIGSEFTT 121
QY 187 IENQPFPAATYRRHSGSVTVVCGSLISPCWV SATHCFIDYKXEDYIVYIGRSPLNS 246
| | | | |
Db 122 IENQPFPAATYRRHSGSVTVVCGSLISPCWV SATHCFIDYKXEDYIVYIGRSPLNS 181

QY 247 NTOGEMKFEVENTILHKOVSADTLAHHNDIALKIRSKERCAQPSRTIOTICLPSPWYND 306
| | | | |
Db 182 NTOGEMKFEVENTILHKOVSADTLAHHNDIALKIRSKERCAQPSRTIOTICLPSPWYND 241
QY 307 POFGTCSEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHYYSSEVTTKMLCAADPQ 366
| | | | |
Db 241 POFGTCSEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHYYSSEVTTKMLCAADPQ 301
QY 367 WKTSCQSGSGGPLVCSLQGRMTLTGIVSWGRGALKDKRQVYTRVSHFLPWIRSHKKEE 426
| | | | |
Db 302 WKTSCQSGSGGPLVCSLQGRMTLTGIVSWGRGALKDKRQVYTRVSHFLPWIRSHKKEE 361
QY 427 NGAL 431
| | | | |
Db 362 NGAL 366

RESULT 15

US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011

GENERAL INFORMATION:

APPLICANT: WENDEL, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOCZA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/967,024C

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P 44 42 665.8

FILING DATE: 30-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 148/42444

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 393 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-967-024C-24

Query Match 84.5%; Score 2022; DB 3; Length 393;

Best Local Similarity 100.0%; Pred. No. 2.9e-166;

Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 SKTCYEGNGHFFRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQGLGKHNYCRND 126
| | | | |
Db 2 SKTCYEGNGHFFRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQGLGKHNYCRND 61
QY 127 NRRPWCYVQVGLKPLVQECMWHDCADGKRPSSPEELKFCQCGQKTLRPRFKIIGSEFTT 186
| | | | |

Db	62	NRRRPWCYVGVGLKPLVQECMHDCADGKKPSSPPEELKFOCGQKTLRPFRKIIIGGEFTT	121
Qy	187	IENQPMFAIYRRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDIYVYLGSRRLNS	246
Db	122	IENQPMFAIYRRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDIYVYLGSRRLNS	181
Qy	247	NTQEMKEFEVENLILHKDYADTLAHNDIALLKIRSKEGRCAPSRITQITCLPSMYND	306
Db	182	NTQEMKEFEVENLILHKDYADTLAHNDIALLKIRSKEGRCAPSRITQITCLPSMYND	241
Qy	307	POFGTSCETGPKENSTDYLPBQLKMTYVKLISHRECOQPHYGSEVTTKMLCAADPQ	366
Db	242	POFGTSCETGPKENSTDYLPBQLKMTYVKLISHRECOQPHYGSEVTTKMLCAADPQ	301
Qy	367	WKTDSQCGDSGSLVCSLQGRMTLTGIVSWGRGCALDKKPGVYTRVSHFLPWIRSHTKEE	426
Db	302	WKTDSQCGDSGSLVCSLQGRMTLTGIVSWGRGCALDKKPGVYTRVSHFLPWIRSHTKEE	361
Qy	427	NGIAL 431	
Db	362	NGIAL 366	

Search completed: March 18, 2004, 11:21:21
 Job time : 50.9847 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:19:25 ; Search time 128.263 Seconds
(without alignments)
870.166 Million cell updates/sec

Title: US-10-076-421-2
Perfect score: 2394
Sequence: 1 MRLRLRLLCVLVSDSKG.....VSHFLPMSHTKENGAL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubppa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubppa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB.pep.*
- 13: /cgn2_6/prodata/2/pubppa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubppa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubppa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2394	100.0	431	12 US-10-411-037-34	Sequence 34, Appl
2	2394	100.0	431	13 US-10-076-421-2	Sequence 2, Appl
3	2394	100.0	431	14 US-10-171-311-184	Sequence 184, Appl
4	2394	100.0	431	14 US-10-301-822-161	Sequence 161, Appl
5	2394	100.0	431	14 US-10-131-985-21	Sequence 21, Appl
6	2394	100.0	431	15 US-10-295-027-414	Sequence 414, Appl
7	2394	100.0	431	15 US-10-295-027-1275	Sequence 1275, Appl
8	2394	99.9	431	14 US-10-247-671-149	Sequence 149, Appl
9	2394	99.9	431	14 US-10-193-656-4	Sequence 4, Appl
10	2376	99.2	431	15 US-09-264-468B-1	Sequence 1, Appl
11	2376	96.8	445	15 US-10-360-101-266	Sequence 266, Appl
12	2301	96.1	411	9 US-09-880-503-3	Sequence 3, Appl
13	2291	95.7	411	15 US-10-407-821-2	Sequence 2, Appl
14	2243	93.7	403	9 US-09-880-503-6	Sequence 6, Appl
15	1737	72.6	323	9 US-09-880-503-7	Sequence 7, Appl

15	1628	68.0	337	14 US-10-106-698-6266	Sequence 6266, Appl
17	1628	68.0	337	15 US-10-264-049-2927	Sequence 2927, Appl
18	1508	63.0	276	9 US-09-880-503-5	Sequence 5, Appl
19	1464	61.2	268	15 US-10-407-821-3	Sequence 3, Appl
20	1333	55.7	246	9 US-09-264-468B-2	Sequence 2, Appl
21	1318	55.1	241	10 US-09-898-837A-47	Sequence 47, Appl
22	1096.5	45.8	650	15 US-10-401-077-1	Sequence 1, Appl
23	868.5	36.3	527	10 US-09-987-457-18	Sequence 18, Appl
24	868.5	36.3	527	10 US-09-987-455-19	Sequence 19, Appl
25	868.5	36.3	527	15 US-10-360-101-203	Sequence 203, Appl
26	868.5	36.3	562	9 US-09-969-271-7	Sequence 7, Appl
27	868.5	36.3	562	9 US-09-974-298-145	Sequence 145, Appl
28	868.5	36.3	562	12 US-10-411-037-26	Sequence 26, Appl
29	868.5	36.3	562	14 US-10-193-656-8	Sequence 8, Appl
30	868.5	36.3	562	14 US-10-443-701-4	Sequence 4, Appl
31	837	35.0	143	9 US-09-880-503-8	Sequence 8, Appl
32	793	33.1	135	9 US-09-880-503-4	Sequence 4, Appl
33	793	33.1	138	9 US-09-984-186-12	Sequence 12, Appl
34	793	33.1	138	14 US-10-237-667-12	Sequence 12, Appl
35	793	33.1	138	14 US-10-237-708-12	Sequence 12, Appl
36	793	33.1	138	14 US-10-237-866-12	Sequence 12, Appl
37	793	33.1	138	14 US-10-237-871-12	Sequence 12, Appl
38	793	33.1	138	14 US-10-237-624-12	Sequence 12, Appl
39	785.5	32.8	372	9 US-09-084-491A-3	Sequence 3, Appl
40	785.5	32.8	372	13 US-10-102-704-3	Sequence 3, Appl
41	780	32.6	354	10 US-09-987-457-10	Sequence 10, Appl
42	780	32.6	354	10 US-09-987-455-11	Sequence 11, Appl
43	780	32.6	377	10 US-09-987-455-8	Sequence 8, Appl
44	741	31.0	343	10 US-09-987-457-14	Sequence 14, Appl
45	741	31.0	343	10 US-09-987-455-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-10-411-037-34
Sequence 34, Application US/10411037
Publication No. US20040043446A1
GENERAL INFORMATION:
APPLICANT: Neose Technologies, Inc.
APPLICANT: Defreese, Shawn
APPLICANT: Zopf, David
APPLICANT: Bayer, Robert
APPLICANT: Hakes, David
APPLICANT: Chen, Xi
APPLICANT: Bove, Caryn
TITLE OF INVENTION: ALPHA GALACTOSIDASE A
FILE REFERENCE: 040853-01-5082
CURRENT APPLICATION NUMBER: US/10/411,037
CURRENT FILING DATE: 2003-04-09
PRIOR APPLICATION NUMBER: US 60/328,523
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/344,692
PRIOR FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: US 60/387,292
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: US 60/391,777
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/396,594
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: US 60/404,249
PRIOR FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: US 60/407,527
PRIOR FILING DATE: 2002-08-28
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-411-037-34

Query Match 100.0%; Score 2394; DB 12; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.1e-200;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLCVLVSDSKSGNELHQPNSNCDCLNGGTGVSNKYFSNIHWCNCPKFKGGQ 60
 DB 1 MRALLARLLCVLVSDSKSGNELHQPNSNCDCLNGGTGVSNKYFSNIHWCNCPKFKGGQ 60
 QY HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 DB HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 QY HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 DB HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 QY YCRNPDNRPRPWCYVYGLKPLVQECMVHDCADGKSSPPEELKFCQGGKTLRPRKII 180
 DB YCRNPDNRPRPWCYVYGLKPLVQECMVHDCADGKSSPPEELKFCQGGKTLRPRKII 180
 QY GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWISATHCFTIDYPKEDYIYVLG 240
 DB GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWISATHCFTIDYPKEDYIYVLG 240
 QY RSRINNTQSGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICL 300
 DB RSRINNTQSGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICL 300
 QY PSMYNDPQFGTSCETITGFGKNSDTLYPEQLKMTVVKLISHRECQPHYGSEVTTKML 360
 DB PSMYNDPQFGTSCETITGFGKNSDTLYPEQLKMTVVKLISHRECQPHYGSEVTTKML 360
 QY CAADPQWKTDSCQDGGGPLVCSLQGRMTLTGIVSWRGALDKDKPGVYTRVSHFLPMIR 420
 DB CAADPQWKTDSCQDGGGPLVCSLQGRMTLTGIVSWRGALDKDKPGVYTRVSHFLPMIR 420
 QY SHTKEENGAL 431
 DB SHTKEENGAL 431

RESULT 2

US-10-076-421-2
 ; Sequence 2, Application US/10076421
 ; Publication No. US20020193304A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WADA, MANABU
 ; APPLICANT: WADA, MAOKO
 ; TITLE OF INVENTION: ANTI-HIV AGENTS
 ; FILE REFERENCE: HAYAK-9
 ; CURRENT APPLICATION NUMBER: US/10/076,421
 ; CURRENT FILING DATE: 2002-05-17
 ; PRIOR APPLICATION NUMBER: JP 2001-42655
 ; PRIOR FILING DATE: 2001-02-20
 ; PRIOR APPLICATION NUMBER: JP 2001-184284
 ; PRIOR FILING DATE: 2001-06-19
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-076-421-2

Query Match 100.0%; Score 2394; DB 13; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.1e-200;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLCVLVSDSKSGNELHQPNSNCDCLNGGTGVSNKYFSNIHWCNCPKFKGGQ 60
 DB 1 MRALLARLLCVLVSDSKSGNELHQPNSNCDCLNGGTGVSNKYFSNIHWCNCPKFKGGQ 60
 QY HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 DB HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 QY HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 DB HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120

QY YCRNPDNRPRPWCYVYGLKPLVQECMVHDCADGKSSPPEELKFCQGGKTLRPRKII 180
 DB YCRNPDNRPRPWCYVYGLKPLVQECMVHDCADGKSSPPEELKFCQGGKTLRPRKII 180
 QY GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWISATHCFTIDYPKEDYIYVLG 240
 DB GGEFTTIENQPMFAIYRRHGGSVTVYVCGSLISPCWISATHCFTIDYPKEDYIYVLG 240
 QY RSRINNTQSGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICL 300
 DB RSRINNTQSGEMKEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICL 300
 QY PSMYNDPQFGTSCETITGFGKNSDTLYPEQLKMTVVKLISHRECQPHYGSEVTTKML 360
 DB PSMYNDPQFGTSCETITGFGKNSDTLYPEQLKMTVVKLISHRECQPHYGSEVTTKML 360
 QY CAADPQWKTDSCQDGGGPLVCSLQGRMTLTGIVSWRGALDKDKPGVYTRVSHFLPMIR 420
 DB CAADPQWKTDSCQDGGGPLVCSLQGRMTLTGIVSWRGALDKDKPGVYTRVSHFLPMIR 420
 QY SHTKEENGAL 431
 DB SHTKEENGAL 431

RESULT 3

US-10-171-311-184
 ; Sequence 184, Application US/10171311
 ; Publication No. US2003008270A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Chen, Yan
 ; APPLICANT: Zhao, Xumei
 ; APPLICANT: Monahan, John
 ; APPLICANT: Kamatkar, Shubhangi
 ; APPLICANT: Glatz, Karen
 ; APPLICANT: Gannavarapu, Manjula
 ; APPLICANT: Hoerish, Sebastian
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
 ; TITLE OF INVENTION: OF CERVICAL CANCER
 ; FILE REFERENCE: MFI-035
 ; CURRENT APPLICATION NUMBER: US/10/171,311
 ; CURRENT FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: US 60/298,159
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/298,155
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/335,936
 ; PRIOR FILING DATE: 2001-11-14
 ; NUMBER OF SEQ ID NOS: 238
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 184
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-171-311-184

Query Match 100.0%; Score 2394; DB 14; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.1e-200;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLCVLVSDSKSGNELHQPNSNCDCLNGGTGVSNKYFSNIHWCNCPKFKGGQ 60
 DB 1 MRALLARLLCVLVSDSKSGNELHQPNSNCDCLNGGTGVSNKYFSNIHWCNCPKFKGGQ 60
 QY HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 DB HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 QY HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 DB HCEIDSKTCYEGNGHFRGKASTDTWGRPCLPWNASATVLOQTYHARSALQGLGKHN 120
 QY YCRNPDNRPRPWCYVYGLKPLVQECMVHDCADGKSSPPEELKFCQGGKTLRPRKII 180
 DB YCRNPDNRPRPWCYVYGLKPLVQECMVHDCADGKSSPPEELKFCQGGKTLRPRKII 180

```

QY 181 GGEFTTIENQWPAIYRRHGGSVTVVCGGSLISPCWVISAHCFIDYPKKEDYIYLG 240
DB 181 GGEFTTIENQWPAIYRRHGGSVTVVCGGSLISPCWVISAHCFIDYPKKEDYIYLG 240
QY 241 RSRNSNTQGEEMKFEVENLILHKDYADTLAHNDIALALKIRSEKGRCAQPSRTIOTICL 300
DB 241 RSRNSNTQGEEMKFEVENLILHKDYADTLAHNDIALALKIRSEKGRCAQPSRTIOTICL 300
QY 301 PSWYNDPOFGTSCETITGFKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSWYNDPOFGTSCETITGFKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPOWKTDSCGDSGGPLVCSLQGRMTLTGIYVSWRGALCKDKPGVTVRVSHPFLMIR 420
DB 361 CAADPOWKTDSCGDSGGPLVCSLQGRMTLTGIYVSWRGALCKDKPGVTVRVSHPFLMIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

```

RESULT 4
US-10-301-822-161

```

; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kametkar, Shubhangt
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MEMO1-029P22NM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: PaetSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

```

Query Match 100.0%; Score 2394; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRALLARLLLCVIVSDSKSGNELHGVPSNCDCLNGTGVSNKYFSNIHMCNCPKKEGGQ 60
DB 1 MRALLARLLLCVIVSDSKSGNELHGVPSNCDCLNGTGVSNKYFSNIHMCNCPKKEGGQ 60
QY 61 HCEIDSKTCYEGNGHFRYRGKASTDTMGRPCLPWNSATVYLQOTYHAHRSALQLGIGKN 120
DB 61 HCEIDSKTCYEGNGHFRYRGKASTDTMGRPCLPWNSATVYLQOTYHAHRSALQLGIGKN 120
QY 121 YCNPNDNRPRPWCYVGVGLKPLVQECVMDCADGKKSPPEELKFCQCGQKTLRPRFKII 180
DB 121 YCNPNDNRPRPWCYVGVGLKPLVQECVMDCADGKKSPPEELKFCQCGQKTLRPRFKII 180
QY 181 GGEFTTIENQWPAIYRRHGGSVTVVCGGSLISPCWVISAHCFIDYPKKEDYIYLG 240
DB 181 GGEFTTIENQWPAIYRRHGGSVTVVCGGSLISPCWVISAHCFIDYPKKEDYIYLG 240

```

```

DB 181 GGEFTTIENQWPAIYRRHGGSVTVVCGGSLISPCWVISAHCFIDYPKKEDYIYLG 240
QY 241 RSRNSNTQGEEMKFEVENLILHKDYADTLAHNDIALALKIRSEKGRCAQPSRTIOTICL 300
DB 241 RSRNSNTQGEEMKFEVENLILHKDYADTLAHNDIALALKIRSEKGRCAQPSRTIOTICL 300
QY 301 PSWYNDPOFGTSCETITGFKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSWYNDPOFGTSCETITGFKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPOWKTDSCGDSGGPLVCSLQGRMTLTGIYVSWRGALCKDKPGVTVRVSHPFLMIR 420
DB 361 CAADPOWKTDSCGDSGGPLVCSLQGRMTLTGIYVSWRGALCKDKPGVTVRVSHPFLMIR 420
QY 421 SHTKEENGIAL 431
DB 421 SHTKEENGIAL 431

```

RESULT 5
US-10-131-985-21

```

; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

```

Query Match 100.0%; Score 2394; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRALLARLLLCVIVSDSKSGNELHGVPSNCDCLNGTGVSNKYFSNIHMCNCPKKEGGQ 60
DB 1 MRALLARLLLCVIVSDSKSGNELHGVPSNCDCLNGTGVSNKYFSNIHMCNCPKKEGGQ 60
QY 61 HCEIDSKTCYEGNGHFRYRGKASTDTMGRPCLPWNSATVYLQOTYHAHRSALQLGIGKN 120
DB 61 HCEIDSKTCYEGNGHFRYRGKASTDTMGRPCLPWNSATVYLQOTYHAHRSALQLGIGKN 120
QY 121 YCNPNDNRPRPWCYVGVGLKPLVQECVMDCADGKKSPPEELKFCQCGQKTLRPRFKII 180
DB 121 YCNPNDNRPRPWCYVGVGLKPLVQECVMDCADGKKSPPEELKFCQCGQKTLRPRFKII 180
QY 181 GGEFTTIENQWPAIYRRHGGSVTVVCGGSLISPCWVISAHCFIDYPKKEDYIYLG 240
DB 181 GGEFTTIENQWPAIYRRHGGSVTVVCGGSLISPCWVISAHCFIDYPKKEDYIYLG 240
QY 241 RSRNSNTQGEEMKFEVENLILHKDYADTLAHNDIALALKIRSEKGRCAQPSRTIOTICL 300
DB 241 RSRNSNTQGEEMKFEVENLILHKDYADTLAHNDIALALKIRSEKGRCAQPSRTIOTICL 300
QY 301 PSWYNDPOFGTSCETITGFKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSWYNDPOFGTSCETITGFKENSTDYLYPEOLKMTYVKLISHRECOQPHYGSEVTTKML 360

```

QY 361 CAADPOWKTDSGCGDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPOWKTDSGCGDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTKEENGLAL 431
DB 421 SHTKEENGLAL 431

RESULT 6
US-10-295-027-414
Sequence 414, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: US/10/295,027
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 414
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-414

Query Match 100.0%; Score 2394; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLATLLLCVLLVSDSKSNNELHQVPSNCDCLNGSTCVSNKXFSNIHMCNCPKKGQ 60
DB 1 MRALLATLLLCVLLVSDSKSNNELHQVPSNCDCLNGSTCVSNKXFSNIHMCNCPKKGQ 60
QY 61 HCEIDSKTCYEGNGHYRGKASTPTMGRPCLPNSATVYQOTYHARSALQGLGKKN 120
DB 61 HCEIDSKTCYEGNGHYRGKASTPTMGRPCLPNSATVYQOTYHARSALQGLGKKN 120
QY 121 YCRNPDRRRPWCYVGLKPLVQECMVHDCADGKSSPEELKFCGQKTLRPRFKII 180
DB 121 YCRNPDRRRPWCYVGLKPLVQECMVHDCADGKSSPEELKFCGQKTLRPRFKII 180

DB 121 YCRNPDRRRPWCYVGLKPLVQECMVHDCADGKSSPEELKFCGQKTLRPRFKII 180
QY 181 GGEFTTLENDPWFALYRRRGSSVTVYCGSSLISPCWYTSATHCFLDPKEDYLVYIG 240
DB 181 GGEFTTLENDPWFALYRRRGSSVTVYCGSSLISPCWYTSATHCFLDPKEDYLVYIG 240
QY 241 RSRINSTQCEMKFEVENLILHROVSADTLAHNDIALKIRSEKRCAPSRRTIOTICL 300
DB 241 RSRINSTQCEMKFEVENLILHROVSADTLAHNDIALKIRSEKRCAPSRRTIOTICL 300
QY 301 PSMTNDPQGTSCETIGFGKENSVDLYLPOLMTVVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSMTNDPQGTSCETIGFGKENSVDLYLPOLMTVVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPOWKTDSGCGDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
DB 361 CAADPOWKTDSGCGDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIR 420
QY 421 SHTKEENGLAL 431
DB 421 SHTKEENGLAL 431

RESULT 7
US-10-295-027-1275
Sequence 1275, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: US/10/295,027
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1275
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-1275

Query Match 100.0%; Score 2394; DB 15; Length 431;

Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVIVVSDSKSNEHQVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFKGGQ 60
DB 1 MRALLARLLLCVIVVSDSKSNEHQVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFKGGQ 60
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
QY 121 YCNPNDPRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
DB 121 YCNPNDPRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 240
QY 241 RSRINSTQEMKEFEVENLIHNDYADTLAHNDIALALKIRSEKRCQAPSRITQITCL 300
DB 241 RSRINSTQEMKEFEVENLIHNDYADTLAHNDIALALKIRSEKRCQAPSRITQITCL 300
QY 301 PSYNDPQFQTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSYNDPQFQTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPQMKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALDKXGVTYRVSHFLPMIR 420
DB 361 CAADPQMKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALDKXGVTYRVSHFLPMIR 420
QY 421 SHTKEENGAL 431
DB 421 SHTKEENGAL 431

RESULT 8

US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ. ID NOS: 186
; SOFTWARE: PERL Program
; SEQ. ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match 99.8%; Score 2392; DB 14; Length 431;
Best Local Similarity 99.8%; Pred. No. 1.6e-200;
Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVIVVSDSKSNEHQVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFKGGQ 60
DB 1 MRALLARLLLCVIVVSDSKSNEHQVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFKGGQ 60
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
QY 121 YCNPNDPRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
DB 121 YCNPNDPRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 240

QY 121 YCNPNDPRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
DB 121 YCNPNDPRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 240
QY 241 RSRINSTQEMKEFEVENLIHNDYADTLAHNDIALALKIRSEKRCQAPSRITQITCL 300
DB 241 RSRINSTQEMKEFEVENLIHNDYADTLAHNDIALALKIRSEKRCQAPSRITQITCL 300
QY 301 PSYNDPQFQTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
DB 301 PSYNDPQFQTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKML 360
QY 361 CAADPQMKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALDKXGVTYRVSHFLPMIR 420
DB 361 CAADPQMKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALDKXGVTYRVSHFLPMIR 420
QY 421 SHTKEENGAL 431
DB 421 SHTKEENGAL 431

RESULT 9

US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAL, Rikard
; APPLICANT: Li, Jinn
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/US77-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ. ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ. ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

Query Match 99.8%; Score 2391; DB 14; Length 431;
Best Local Similarity 99.8%; Pred. No. 1.9e-200;
Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVIVVSDSKSNEHQVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFKGGQ 60
DB 1 MRALLARLLLCVIVVSDSKSNEHQVPSNCDCLNGSTCVSNKYFSNIHMCNCPKFKGGQ 60
QY 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
DB 61 HCEIDSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVILQOTYHAHRSALQIGLKEN 120
QY 121 YCNPNDPRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
DB 121 YCNPNDPRRPWCYVQGLKPLVQECMVHDCADGKSSPPEELKFCQCGKTLRPRFKII 180
QY 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 240
DB 181 GGEFTTIENQPMFAIYRRHRGGSVTYVCGSLISPCWVISAHCFIDYPKKEDYIYVLG 240

Db 181 GGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLMSPCWVISAHCHFDYPKKEDYIYVIG 240
QY 241 RSRINSTOGEKMEKEVENLILHKYSDATLAHNDIALIKIRKEGRCAQPSRTIOTICL 300
Db 241 RSRINSTOGEKMEKEVENLILHKYSDATLAHNDIALIKIRKEGRCAQPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTTKML 360
Db 301 PSMYNDPOFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTTKML 360
QY 361 CAADPOWKTDSCQDGSGLVCSLQGRMTLTGIVSWGRCALKDXKPGVYTRVSHFLPWIR 420
Db 361 CAADPOWKTDSCQDGSGLVCSLQGRMTLTGIVSWGRCALKDXKPGVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
Db 421 SHTKEENGIAL 431

RESULT 10

US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310. US. P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match 99.2%; Score 2376; DB 9; Length 431;
Best Local Similarity 99.5%; Pred. No. 3.9e-199;
Matches 429; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRALLARLLLCVLVSDSKSNEIHOVPSNCDCLNGGTCVSNKYF--SNIMWCNCPKFKG 60
Db 1 MRALLARLLLCVLVSDSKSNEIHOVPSNCDCLNGGTCVSNKYF--SNIMWCNCPKFKG 60
QY 61 HCEIDSKTCYEGNGHFGYRGASTDYMGKPCLPWNSATVTLQOTYHAHRSALQGLGKH 120
Db 61 HCEIDSKTCYEGNGHFGYRGASTDYMGKPCLPWNSATVTLQOTYHAHRSALQGLGKH 120
QY 121 YCRNPDRRRRWCYVQVGLKPLVQECMVDCAADGKSPPEBELKFGCGOKTLARPKKI 180
Db 121 YCRNPDRRRRWCYVQVGLKPLVQECMVDCAADGKSPPEBELKFGCGOKTLARPKKI 180

QY 181 GGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLISPWCVISAHCHFDYPKKEDYIYVIG 240
Db 181 GGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLISPWCVISAHCHFDYPKKEDYIYVIG 240
QY 241 RSRINSTOGEKMEKEVENLILHKYSDATLAHNDIALIKIRKEGRCAQPSRTIOTICL 300
Db 241 RSRINSTOGEKMEKEVENLILHKYSDATLAHNDIALIKIRKEGRCAQPSRTIOTICL 300
QY 301 PSMYNDPOFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTTKML 360
Db 301 PSMYNDPOFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTTKML 360
QY 361 CAADPOWKTDSCQDGSGLVCSLQGRMTLTGIVSWGRCALKDXKPGVYTRVSHFLPWIR 420
Db 361 CAADPOWKTDSCQDGSGLVCSLQGRMTLTGIVSWGRCALKDXKPGVYTRVSHFLPWIR 420
QY 421 SHTKEENGIAL 431
Db 421 SHTKEENGIAL 431

RESULT 11

US-10-360-101-266
; Sequence 266, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly) peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 266
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of urokinase
US-10-360-101-266

Query Match 96.8%; Score 2317; DB 15; Length 445;
Best Local Similarity 96.9%; Pred. No. 5.9e-194;
Matches 431; Conservative 0; Mismatches 0; Indels 14; Gaps 7;

QY 1 MRALLARLLLCVLVSDSKSNEIHOVPSNCDCLNGGTCVSNKYF--SNIMWCNCPKFKG 58
Db 1 MRALLARLLLCVLVSDSKSNEIHOVPSNCDCLNGGTCVSNKYF--SNIMWCNCPKFKG 60
QY 59 GCHCEIDSKTCYEGNGHFGYRGASTDYMGKPCLPWNSATVTLQOTY--HAHRSALQGL 116
Db 61 GCHCEIDSKTCYEGNGHFGYRGASTDYMGKPCLPWNSATVTLQOTY--HAHRSALQGL 120
QY 117 GKNYCRNPDRRRRWCYVQVGLKPLVQECMVDCAADGKSPPEBELKFGCGOKTLAR 174
Db 121 GKNYCRNPDRRRRWCYVQVGLKPLVQECMVDCAADGKSPPEBELKFGCGOKTLAR 180
QY 175 PREKIIIGGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLISPWCVISA--THCFIDYPKK 232
Db 181 PREKIIIGGEFTTIENQPMWFAIYRRHRGGSVTYVCGSLISPWCVISA--THCFIDYPKK 240
QY 233 EDYIVYIGSRRLNSNTGGEKMEKEVENLILHKYSDATLAHNDIALIKIR--KEGRCAQ 230
Db 241 EDYIVYIGSRRLNSNTGGEKMEKEVENLILHKYSDATLAHNDIALIKIR--KEGRCAQ 300
QY 291 PSRTIOTICLPSMYNDPOFGTSCETITGFGKENSTDYLYPEQLKMTVVKL--SHRECOOP 348
Db 301 PSRTIOTICLPSMYNDPOFGTSCETITGFGKENSTDYLYPEQLKMTVVKL--SHRECOOP 360

QY 349 HYGSEVTTKMLCAADPQWKTDSCQDSGGGLVCSLQGRMTLTGIVSWGRG--CALDKP 406
DB 361 HYGSEVTTKMLCAADPQWKTDSCQDSGGGLVCSLQGRMTLTGIVSWGRGFCALDKP 420
QY 407 GYTVRSHFLPMIRSHKTEENGAL 431
DB 421 GYTVRSHFLPMIRSHKTEENGAL 445

RESULT 12

US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match 96.1%; Score 2301; DB 9; Length 411;
Best Local Similarity 100.0%; Pred. No. 1,3e-192;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SNEHQVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFGQHCIDSKTCYEGNGHYRG 80
DB 1 SNEHQVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFGQHCIDSKTCYEGNGHYRG 60
QY 81 KASTDMGRPCLPMNSATVLOQTYHAHRSALQGLGKANYCNPDRRRPWCYVGYLK 140
DB 61 KASTDMGRPCLPMNSATVLOQTYHAHRSALQGLGKANYCNPDRRRPWCYVGYLK 120
QY 141 PLVQECMVHDCAGKRPSPPEELKFCQCGKTLRPFKIIGSEFTTENPWFALYRRH 200
DB 121 PLVQECMVHDCAGKRPSPPEELKFCQCGKTLRPFKIIGSEFTTENPWFALYRRH 180
QY 201 RGSSTVTVCGSLISPCWVISAHCFLIDPKEDYIVYLSRSLNSNTQEMKFEVENLI 260
DB 181 RGSSTVTVCGSLISPCWVISAHCFLIDPKEDYIVYLSRSLNSNTQEMKFEVENLI 240
QY 261 LKHDYADTLAHNDIALKIRSEKRCAPSRITQICLPSMYNDPQFSGCEITGFGK 320
DB 241 LKHDYADTLAHNDIALKIRSEKRCAPSRITQICLPSMYNDPQFSGCEITGFGK 300
QY 321 ENSTDVLYPEQLKMTVVKLISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPL 380
DB 301 ENSTDVLYPEQLKMTVVKLISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPL 360
QY 381 VCSLQGRMTLTGIVSWGRGALDKDKPGVYTVRSHFLPMIRSHKTEENGAL 431
DB 361 VCSLQGRMTLTGIVSWGRGALDKDKPGVYTVRSHFLPMIRSHKTEENGAL 411

RESULT 13

US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS

; FILE REFERENCE: UTSN:02205
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match 95.7%; Score 2291; DB 15; Length 411;
Best Local Similarity 99.8%; Pred. No. 1e-191;
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 SNEHQVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFGQHCIDSKTCYEGNGHYRG 80
DB 1 SNEHQVPSNCDCLNGSTCVSNKYFSNIHWCNCPKFGQHCIDSKTCYEGNGHYRG 60
QY 81 KASTDMGRPCLPMNSATVLOQTYHAHRSALQGLGKANYCNPDRRRPWCYVGYLK 140
DB 61 KASTDMGRPCLPMNSATVLOQTYHAHRSALQGLGKANYCNPDRRRPWCYVGYLK 120
QY 141 PLVQECMVHDCAGKRPSPPEELKFCQCGKTLRPFKIIGSEFTTENPWFALYRRH 200
DB 121 PLVQECMVHDCAGKRPSPPEELKFCQCGKTLRPFKIIGSEFTTENPWFALYRRH 180
QY 201 RGSSTVTVCGSLISPCWVISAHCFLIDPKEDYIVYLSRSLNSNTQEMKFEVENLI 260
DB 181 RGSSTVTVCGSLISPCWVISAHCFLIDPKEDYIVYLSRSLNSNTQEMKFEVENLI 240
QY 261 LKHDYADTLAHNDIALKIRSEKRCAPSRITQICLPSMYNDPQFSGCEITGFGK 320
DB 241 LKHDYADTLAHNDIALKIRSEKRCAPSRITQICLPSMYNDPQFSGCEITGFGK 300
QY 321 ENSTDVLYPEQLKMTVVKLISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPL 380
DB 301 ENSTDVLYPEQLKMTVVKLISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPL 360
QY 381 VCSLQGRMTLTGIVSWGRGALDKDKPGVYTVRSHFLPMIRSHKTEENGAL 431
DB 361 VCSLQGRMTLTGIVSWGRGALDKDKPGVYTVRSHFLPMIRSHKTEENGAL 411

RESULT 14

US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match 93.7%; Score 2243; DB 9; Length 403;
Best Local Similarity 98.1%; Pred. No. 1.5e-187;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```

QY 21 SNEHQVPSNCDCLNGGTGCVSNKYFSNTHMNCNCPKKGCGHCEIDKSKTCYEGNGHFRG 80
Db 1 SNEHQVPSNCDCLNGGTGCVSNKYFSNTHMNCNCPKKGCGHCEIDKSKTCYEGNGHFRG 60
QY 81 KASDTMGRPCLPNNSATVLOQTYHAHRSDALQGLGKHNCRNPDNRNRPWCYVQVGLK 140
Db 61 KASDTMGRPCLPNNSATVLOQTYHAHRSDALQGLGKHNCRNPDNRNRPWCYVQVGLK 120
QY 141 PLYOECWHDCAADGKPSPEELKFGCGKTLRPRFKTIGSEFTTIENQWPAIYRRH 200
Db 121 PLYOECWHDCAADGK-----LKFQCGQKTLRPRFKTIGSEFTTIENQWPAIYRRH 172
QY 201 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVYIGRSRLNSNTQGMKEVENTLI 260
Db 173 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVYIGRSRLNSNTQGMKEVENTLI 232
QY 261 LHKDYSADTLAHNDIALALKIRSKRGRCAPSRITQITCLPSMTNDPQFTSCETITGFGK 320
Db 233 LHKDYSADTLAHNDIALALKIRSKRGRCAPSRITQITCLPSMTNDPQFTSCETITGFGK 292
QY 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCGDSGGPL 380
Db 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCGDSGGPL 352
QY 381 VCSLQGRMTLTGIVSMRGCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 431
Db 353 VCSLQGRMTLTGIVSMRGCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 403

```

RESULT 15

```

US-09-880-503-7
; Sequence 7, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212, 847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-7

```

```

Query Match 72.6%; Score 1737; DB 9; Length 323;
Best Local Similarity 78.6%; Pred. No. 1.9e-143;
Matches 323; Conservative 0; Mismatches 0; Indels 88; Gaps 1;

```

```

QY 21 SNEHQVPSNCDCLNGGTGCVSNKYFSNTHMNCNCPKKGCGHCEIDKSKTCYEGNGHFRG 80
Db 1 SNEHQVPSNCDCLNGGTGCVSNKYFSNTHMNCNCPKKGCGHCEIDKSKTCYEGNGHFRG 47
QY 81 KASDTMGRPCLPNNSATVLOQTYHAHRSDALQGLGKHNCRNPDNRNRPWCYVQVGLK 140
Db 48 ----- 47
QY 141 PLYOECWHDCAADGKPSPEELKFGCGKTLRPRFKTIGSEFTTIENQWPAIYRRH 200
Db 48 -----KPSPEELKFGCGKTLRPRFKTIGSEFTTIENQWPAIYRRH 92
QY 201 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVYIGRSRLNSNTQGMKEVENTLI 260
Db 93 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVYIGRSRLNSNTQGMKEVENTLI 152
QY 261 LHKDYSADTLAHNDIALALKIRSKRGRCAPSRITQITCLPSMTNDPQFTSCETITGFGK 320

```

```

Db 153 LHKDYSADTLAHNDIALALKIRSKRGRCAPSRITQITCLPSMTNDPQFTSCETITGFGK 212
QY 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCGDSGGPL 380
Db 213 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCGDSGGPL 272
QY 381 VCSLQGRMTLTGIVSMRGCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 431
Db 273 VCSLQGRMTLTGIVSMRGCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 323

```

```

Search completed: March 18, 2004, 11:36:07
Job time : 129.513 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:10:10 ; Search time 44.326 Seconds
(without alignments)
935.309 Million cell updates/sec

Title: US-10-076-421-2

Perfect score: 2394

Sequence: 1 MRALLARLLLCVLVSDSKG.....VSHFLPWRSHRTKENGAL 431

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

1: PIR 78: *
2: p1r1: *
3: p1r2: *
4: p1r3: *
5: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2391	99.9	431	1 UKHU	u-plasminogen acti
2	2228	93.1	433	1 UKBAY	u-plasminogen acti
3	1928.5	80.6	442	1 UKPG	u-plasminogen acti
4	1835	76.6	433	1 JN0560	u-plasminogen acti
5	1728.5	72.2	432	1 S18932	u-plasminogen acti
6	1700	71.0	433	1 UKMS	u-plasminogen acti
7	1009	42.1	434	1 A35005	u-plasminogen acti
8	868.5	36.3	477	2 JS0598	t-plasminogen acti
9	868.5	36.3	562	1 UKHUT	t-plasminogen acti
10	867.5	36.2	431	2 JS0599	t-plasminogen acti
11	867.5	36.2	477	1 A34369	t-plasminogen acti
12	847.5	35.4	559	1 A35029	t-plasminogen acti
13	841.5	35.2	477	2 JS0597	t-plasminogen acti
14	833.5	34.8	559	1 A29941	t-plasminogen acti
15	752	31.4	394	2 JS0600	hepatocyte growth
16	735.5	30.7	655	1 A46688	hepatocyte growth
17	718.5	30.0	603	2 S28941	coagulation factor
18	692	28.9	615	1 KEHUI2	coagulation factor
19	674.5	28.2	558	2 JCS478	plasma hyaluronan
20	661.5	27.6	560	1 JCS479	plasma hyaluronan
21	643.5	26.9	553	2 S45281	coagulation factor
22	507.5	21.2	460	2 B61545	plasmin (EC 3.4.21
23	504.5	21.1	790	1 PLUG	plasmin (EC 3.4.21
24	501	20.9	810	1 PLUG	plasmin (EC 3.4.21
25	497	20.8	810	2 B30848	plasmin (EC 3.4.21
26	484.5	20.2	812	1 PLUG	plasmin (EC 3.4.21
27	484	20.2	1420	2 A32869	apoptoprotein(a)
28	474	19.8	4548	1 S00657	apoptoprotein(a) (EC

30	457.5	19.1	455	2 A61545	plasmin (EC 3.4.21
31	445	18.6	761	2 JCS759	brain-specific ser
32	435	18.2	810	2 I46260	plasmin (EC 3.4.21
33	410	17.1	638	1 K0MSPL	prostate kallikrein
34	408.5	17.1	343	1 A57014	prostatein (EC 3.4.
35	406	17.0	711	1 A47136	macrophage-stimula
36	402.5	16.8	417	1 S00845	hepsin (EC 3.4.21.
37	402	16.8	638	1 K0HUP	plasma kallikrein
38	401	16.8	416	1 S33777	hepsin (EC 3.4.21.
39	401	16.8	855	2 JCS731	membrane-bound arg
40	395.5	16.5	248	2 S55066	trypsin (EC 3.4.21
41	394.5	16.5	229	1 TRB0TR	trypsin (EC 3.4.21
42	394.5	16.5	716	1 JCS061	macrophage-stimula
43	393	16.4	728	1 A40332	hepatocyte growth
44	392	16.4	716	1 A40332	macrophage-stimula
45	390.5	16.3	247	2 S13813	trypsin (EC 3.4.21

ALIGNMENTS

RESULT 1

UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N;Contents: urokinase-type plasminogen activator chain A; urokinase-type plasminogen acti
in form

C;Species: Homo sapiens (man)

C;Date: 17-Dec-1982 #sequence_revision 04-Dec-1986 #text change 15-Sep-2000

C;Accession: A00931; 152209; J01012; A37561; 138102; S65783; A37562; A37563; A37564; A356

R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boati, S.; Blasi, F.

Nucleic Acids Res. 13, 2759-2771, 1985

A;Title: The human urokinase-plasminogen activator gene and its promoter.

A;Reference number: A00931; M01D:85215647; PMID:2987867

A;Accession: A00931

A;Molecule type: DNA

A;Residues: 1-431 <RC>

A;Cross-references: GB:K03429; NID:937601; PIDN:CAA26268.1; PID:91834524

A;Note: the authors translated the codon ATG for residue 214 as Ile

R;Nagamine, Y.; Pearson, D.; Gratian, M.

Biochem. Biophys. Res. Commun. 135, 563-569, 1985

A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine

A;Reference number: 152209; M01D:86050639; PMID:3933505

A;Accession: 152209

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 145-161 <NAG1>

A;Cross-references: GB:K03427; NID:9340174; PIDN:AAA61257.1; PID:9340175

R;Nagai, M.; Hiratake, R.; Kaneda, T.; Hayasaka, N.; Aritama, H.; Nishida, M.; Suyama, J

Gene 36, 183-188, 1985

A;Title: Molecular cloning of cDNA coding for human prepro-urokinase.

A;Reference number: J01012; M01D:86056954; PMID:2415429

A;Accession: J01012

A;Molecule type: mRNA

A;Residues: 1-213, 'I', 215-431 <NAG2>

A;Cross-references: GB:K03426; NID:9340155; PIDN:AA971138.1; PID:9340158; GB:D00244; NID

R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.

Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984

A;Title: Identification and primary sequence of an unspliced human urokinase poly(A)+ RN

A;Reference number: A37561; M01D:84272706; PMID:6589620

A;Accession: A37561

A;Molecule type: mRNA

A;Residues: 66-431 <VER>

A;Cross-references: GB:D00244; NID:9220138

R;Jacobs, P.; Crivador, A.; Lortia, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen,

DNA 4, 139-146, 1985

A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human prepr

A;Reference number: 138102; M01D:85203359; PMID:3888571

A;Accession: 138102

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-150, 'W', 152-213, 'T', 215-385, 'C', 387-429, 'V', 431 <JAC>

A;Cross-references: EMBL:X02760; NID:935297; PIDN:CAA6535.1; PID:935298

R; Yoshimoto, M.; Uehiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, K.; Blochm. Biophys. Acta 1293, 83-89, 1996
A:Title: Characterization of single chain urokinase-type plasminogen activator with a no
A:Reference number: 965783; MUID:96186279; PMID:8652631
A:Accession: S65783
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 21-140, 'L', 142-213, 'T', 215-431 <YOS>
A:Cross-references: EMBL:D1143; NID:91311467; PIDN:BA01919.1; PID:g119928
R; Gunzler, W.A.; Steffens, G.J.; Oetting, F.; Kilm, S.M.A.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
A:Title: The primary structure of high molecular mass urokinase from human urine.
A:Reference number: A37562; MUID:83055084; PMID:6754569
A:Accession: A37562
A:Molecule type: protein
A:Residues: 21-177 <GUN>
R; Schaller, U.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergler, W.; Studer, R.O.
Eur. J. Biochem. 125, 251-257, 1982
A:Title: Human low-molecular-weight urinary urokinase. Partial characterization and prel
A:Reference number: A37563; MUID:83003608; PMID:6749491
A:Accession: A37563
A:Molecule type: protein
A:Residues: 156-176; 179-193, 'T', 195, 'T', 197-224 <SCH>
R; Steffens, G.J.; Gunzler, W.A.; Oetting, F.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
A:Title: The complete amino acid sequence of low molecular mass urokinase from human ur
A:Reference number: A37564; MUID:83055099; PMID:6754572
A:Accession: A37564
A:Molecule type: protein
A:Residues: 158-410 <STE>
R; Kettner, E.J.; Buko, A.; Menon, G.; Sarin, V.K.
Biochem. Biophys. Res. Commun. 171, 401-406, 1990
A:Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinat
A:Reference number: A35689; MUID:90365737; PMID:2393398
A:Accession: A35689
A:Molecule type: protein
A:Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
A:Note: identification of a fucose and attempt to determine its attachment site
R; Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazur, A.; Henkin, J.; Goltz
Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
A:Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
A:Reference number: A36697; MUID:91097529; PMID:2125213
A:Accession: A36697
A:Molecule type: protein
A:Residues: 21-34 <RAB>
R; Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
submitted to the Brookhaven Protein Data Bank, July 1993
A:Reference number: A51255; PDB:1KDU
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R; Li, X.; Smith, R.A.G.; Dobson, C.M.
Biochemistry 31, 9562-9571, 1992
A:Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain f
A:Reference number: A44375; MUID:93003110; PMID:1327118
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
R; Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazur, A.P.; Olejniczak,
submitted to the Brookhaven Protein Data Bank, January 1994
A:Reference number: A66822; PDB:1URK
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R; Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
submitted to the Brookhaven Protein Data Bank, July 1995
A:Reference number: A66058; PDB:1LWM
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
A:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
C:Genetics: GDB: PLAU
A:Gene: GDB: PLAU
A:Cross-references: GDB:119497; OMIM:191840
A:Map position: 10q24-10q24
A:Introns: 19/3, 25/1, 65/1, 123/2, 154/1, 227/2, 277/1, 324/1, 373/3
C:Function: proteolytically activates plasminogen
A:Description: proteolytically activates plasminogen
A:Pathway: fibrinolysis
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteinat
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-43/Product: urokinase-type plasminogen activator, single chain form #status predict
F:21-177/Product: urokinase-type plasminogen activator chain A #status experimental <MPA>
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KR>
F:156-177/Product: urokinase-type plasminogen activator chain B #status experimental <MPI
F:179-431/Product: urokinase-type plasminogen activator chain B #status experimental <MPI
F:179-419/Domain: trypsin homology <TRY>
F:31-39, 33-51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372-400
F:38/Binding site: carbohydrate (Thr) (covalent) #status predicted
F:178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
F:224, 275, 376/Active site: His, Asp, Ser #status experimental
F:322/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 99.8%; Score 2391; DB 1; Length 431;
Best Local Similarity 99.8%; Pred. No. 2, 3e-175;
Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALLARLLCVLVVSDSKGNSNELHQVPSNCDLNGTGVSNKYPSNHYMCNCPKFGQ 60
Db 1 MRALLARLLCVLVVSDSKGNSNELHQVPSNCDLNGTGVSNKYPSNHYMCNCPKFGQ 60
Qy HCEIDSKTCYEGNGHFRYKASTDTPMGRPCPMNSATYLCQYHNRSDALQLGKXN 120
Db HCEIDSKTCYEGNGHFRYKASTDTPMGRPCPMNSATYLCQYHNRSDALQLGKXN 120
Qy 61 HCEIDSKTCYEGNGHFRYKASTDTPMGRPCPMNSATYLCQYHNRSDALQLGKXN 120
Db 61 HCEIDSKTCYEGNGHFRYKASTDTPMGRPCPMNSATYLCQYHNRSDALQLGKXN 120
Qy 121 YCNPNRRPWCYVGVGLPLVQECVMDCADGKRPSPBELKFCQCKTLRPFKII 180
Db 121 YCNPNRRPWCYVGVGLPLVQECVMDCADGKRPSPBELKFCQCKTLRPFKII 180
Qy 121 YCNPNRRPWCYVGVGLPLVQECVMDCADGKRPSPBELKFCQCKTLRPFKII 180
Db 121 YCNPNRRPWCYVGVGLPLVQECVMDCADGKRPSPBELKFCQCKTLRPFKII 180
Qy 181 GGEFTTIEQNPFAIYRRHGGSVTVVCGGSLSPCWYISAHNCTIDPKEDYIVYG 240
Db 181 GGEFTTIEQNPFAIYRRHGGSVTVVCGGSLSPCWYISAHNCTIDPKEDYIVYG 240
Qy 181 GGEFTTIEQNPFAIYRRHGGSVTVVCGGSLSPCWYISAHNCTIDPKEDYIVYG 240
Db 181 GGEFTTIEQNPFAIYRRHGGSVTVVCGGSLSPCWYISAHNCTIDPKEDYIVYG 240
Qy 241 RSLNSTQCEMKEFEVNLILHVDYADTLAHNDLALKIRKEKRCAPSTTIQICL 300
Db 241 RSLNSTQCEMKEFEVNLILHVDYADTLAHNDLALKIRKEKRCAPSTTIQICL 300
Qy 241 RSLNSTQCEMKEFEVNLILHVDYADTLAHNDLALKIRKEKRCAPSTTIQICL 300
Db 241 RSLNSTQCEMKEFEVNLILHVDYADTLAHNDLALKIRKEKRCAPSTTIQICL 300
Qy 301 PSMYNDPQFQTSCEITFGKENSTDYLYPEQLMTYVKKLISHRECOQPHYGSEVTTKML 360
Db 301 PSMYNDPQFQTSCEITFGKENSTDYLYPEQLMTYVKKLISHRECOQPHYGSEVTTKML 360
Qy 301 PSMYNDPQFQTSCEITFGKENSTDYLYPEQLMTYVKKLISHRECOQPHYGSEVTTKML 360
Db 301 PSMYNDPQFQTSCEITFGKENSTDYLYPEQLMTYVKKLISHRECOQPHYGSEVTTKML 360
Qy 361 CAADPQWKTDSCGDSGGLVCSLOGRMTLTVISWGRGALKDKPQVTVRVSHFLPMIR 420
Db 361 CAADPQWKTDSCGDSGGLVCSLOGRMTLTVISWGRGALKDKPQVTVRVSHFLPMIR 420
Qy 421 SHKEENGIAL 431
Db 421 SHKEENGIAL 431
Qy 421 SHKEENGIAL 431
Db 421 SHKEENGIAL 431
RESULT 2
UKRAY
u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C:Date: 31-Dec-1991 #sequence, revision 31-Dec-1991 #text, change 18-Jun-1999
C:Accession: S14687; S08651
R; Au, Y.P.T.; Wang, T.W.; Clowes, A.W.
Nucleic Acids Res. 18, 3411, 1990
A:Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen
A:Reference number: S14687; MUID:90287734; PMID:2113276
A:Accession: S14687
A:Molecule type: mRNA
A:Residues: 1-433 <AYU>
A:Cross-references: EMBL:X51935; NID:938130; PIDN:CA36200.1; PID:g38131
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-176/Product: plasminogen activator chain A #status predicted <ACH>
F:30-61/Domain: EGF homology <EGF>
F:169-150/Domain: kringle homology <KR>
F:178-433/Product: plasminogen activator chain B #status predicted <BGH>

F:178-421/Domain: trypsin homology <TRY>
 F:167-298,208-224,216-287,315-384,374-402/Disulfide bonds: #status predicted
 F:223,274,378/Active site: His, Asp, Ser #status predicted
 F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.1%; Score 2228; DB 1; Length 433;
 Best Local Similarity 92.4%; Pred. No. 6,9e-153;
 Matches 401; Conservative 17; Mismatches 12; Indels 4; Gaps 2;

```

QY 1 MRLLARLLLCVIVSDSKSGNELHOVPSNCDCLNGGTCVSNKYSNIHMCNCPKPKFG 60
DB 1 MRLLARLLLCVIVSDSKSGNELHOVPSNCDCLNGGTCVSNKYSNIHMCNCPKPKFG 59
QY 61 HCEIDSKSKTCYEGNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHSDALQGLGKN 120
DB 60 HCEIDSKSKTCYEGNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHSDALQGLGKN 119
QY 121 YCRPNRNRPRPCVYQVGLKPLVOECWVHCDAGKSSPPEELKFCGQKTLRPRKXIT 180
DB 120 YCRPNRNRPRPCVYQVGLKPLVOECWVHCDAGKSSPPEELKFCGQKTLRPRKXIV 179
QY 181 GGEFTTIENQPMFPAIYRRHSGSVTVVCGSLISPCWVISATHCFIDYRKEDYIYVLG 240
DB 180 GGEFTTIENQPMFPAIYRRHSGSVTVVCGSLISPCWVISATHCFIDYRKEDYIYVLG 239
QY 241 RSLNSNTQGEKFEVENILHKDYSADTLAHNDIALKIRSKERCAOPSRITQITCL 300
DB 240 RSLNSNTQGEKFEVENILHEDYSADTLAHNDIALKIRSKERCAOPSRITQITCL 299
QY 301 PSNYNDPO--FGTSCITTFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTT 357
DB 300 PSNYNDPO--FGTSCITTFGKENSTDYLYPEQLKMTVVKLISHRECOOPHYGSEVTT 359
QY 358 KMLCAADPOWKTDSCQDSGGLVCSLQGRMTLTGIYSWGRGALKDKPQVTVRVSHFLP 417
DB 360 KMLCAADPOWKTDSCQDSGGLVCSLQGRMTLTGIYSWGRGALKDKPQVTVRVSHFLP 419
QY 418 WIRSHTEENGIAL 431
DB 420 WIRSHTEENGIAL 433

```

RESULT 3

UNPG u-plasminogen activator (EC 3.4.21.73) precursor - pig

N:Alternate names: uPA
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C:Accession: A00932

R:Nagamine, Y.; Pearson, D.; Altun, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A>Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A:Reference number: A00932; M01D:85087954; PMID:6096832
 A:Accession: A00932

A:Molecule type: DNA
 A:Residues: 1-240, 'H', 242-442 <NA61>
 A:Experimental source: kidney cell line LLC-PK1
 R:Nagamine, Y.
 submitted to the Protein Sequence Database, December 1986

A:Reference number: A3756
 A:Contents annotation: correction to residue 241
 C:Gene(s):

A:Insertions: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:23-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KR>
 F:190-430/Domain: trypsin homology <TRY>
 F:190-430/Domain: trypsin homology <TRY>
 F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted <BGH>
 F:179-310,220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted

F:235,286,387/Active site: His, Asp, Ser #status predicted
 Query Match 80.6%; Score 1928.5; DB 1; Length 442;
 Best Local Similarity 79.4%; Pred. No. 5,8e-140;
 Matches 350; Conservative 33; Mismatches 47; Indels 11; Gaps 2;

```

QY 1 MRLLARLLLCVIVSDSKSGNELHOV--PSNCDCLNGGTCVSNKYSNIHMCNCPKPKFG 58
DB 1 MRLLARLLLCVIVSDSKSGNELHOV--PSNCDCLNGGTCVSNKYSNIHMCNCPKPKFG 60
QY 59 GCHCEIDSKSKTCYEGNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHSDALQGLGK 118
DB 61 GCHCEIDSKSKTCYEGNGHFRGKASTDTMGRPCLPMNSATVLOQTYHAHSDALQGLGK 120
QY 119 HNYCRPNRNRPRPCVYQVGLKPLVOECWVHCDAGKSSPPEELKFCGQKTLRPRKXIT 169
DB 121 HNYCRPNRNRPRPCVYQVGLKPLVOECWVHCDAGKSSPPEELKFCGQKTLRPRKXIV 180
QY 170 OKTLRPRKXITGGEFTTIENQPMFPAIYRRHSGSVTVVCGSLISPCWVISATHCFIDY 229
DB 181 OKTLRPRKXIVGGEFTTIENQPMFPAIYRRHSGSVTVVCGSLISPCWVISATHCFIDY 240
QY 230 PKEDYIYVIGRSLNSNTQGEKFEVENILHKDYSADTLAHNDIALKIRSKERCA 289
DB 241 OKEDYIYVIGRSLNSNTQGEKFEVENILHEDYSADTLAHNDIALKIRSKERCA 300
QY 290 QPSRTIOTICLPYNDPOFGTSCITTFGKENSTDYLYPEQLKMTVVKLISHRECOOPH 349
DB 301 QPSRSIOTICLPYNDPOFGTSCITTFGKENSTDYLYPEQLKMTVVKLISHRECOOPH 360
QY 350 YGSEVTTKMLCAADPOWKTDSCQDSGGLVCSLQGRMTLTGIYSWGRGALKDKPQVTV 409
DB 361 YGSEVTTKMLCAADPOWKTDSCQDSGGLVCSLQGRMTLTGIYSWGRGALKDKPQVTV 420
QY 410 TRYSHFLPWRSHTEENGIAL 430
DB 421 TRYSHFLPWRSHTEENGIAL 441

```

RESULT 4

UN0560 u-plasminogen activator (EC 3.4.21.73) precursor - bovine

N:Alternate names: uPA
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C:Accession: UN0560

R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schlemming, W.D.
 Gene 125, 177-183, 1993
 A>Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induc
 A:Reference number: UN0560; M01D:93216119; PMID:8385052
 A:Accession: UN0560

A:Molecule type: mRNA
 A:Residues: 1-433 <KRA>
 A:Cross-references: GB:U03546; NID:G163800; PIDN:AA51419.1; PID:G163801
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-179/Product: plasminogen activator chain A #status predicted <MA1>
 F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <MA2>
 F:33-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KR>
 F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
 F:181-421/Domain: trypsin homology <TRY>
 F:170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 76.6%; Score 1835; DB 1; Length 433;
 Best Local Similarity 75.1%; Pred. No. 8e-133;
 Matches 325; Conservative 46; Mismatches 60; Indels 2; Gaps 1;

```

QY 1 MRLLARLLLCVIVSDSKSGNELHOV--PSNCDCLNGGTCVSNKYSNIHMCNCPKPKFG 58
DB 1 MRLLARLLLCVIVSDSKSGNELHOV--PSNCDCLNGGTCVSNKYSNIHMCNCPKPKFG 60

```

QY 59 GQCEIDSKTCYGNHGFYRGKASTDTMGRPCLPMNSATVLOQTTHAHRSDALQLGK 118
 Db 61 GEHEIDISTKTCYGNHGFYRGKASTDTMGRPCLPMNSATVLOQTTHAHRSDALQLGK 120
 QY 119 HNYCRNPDRRPMPCYVQVGLKPLVOCMHDCAQDKKPSPPPELKFQCGQKTLRPRFK 178
 Db 121 HNYCRNPDRRPMPCYVQVGLKPLVOCMHDCAQDKKPSPPPELKFQCGQKTLRPRFK 180
 QY 179 IIGEFITIEQPMFAIYRRHGGSYVYCGGSLISPCWVISAATHGIDYPKKEDYIVY 238
 Db 181 IVGGQVTAENQPMFAIYRRHGGSYVYCGGSLISPCWVISAATHGIDYPKKEDYIVY 240
 QY 239 LGRSLNSNTQGEKFEVENLILHKDYSADTLAHNDIALIKIRSKRGCAQPSRTIQT 298
 Db 241 LGQSLNSNTQGEKFEVENLILHKDYSADTLAHNDIALIKIRSKRGCAQPSRTIQT 300
 QY 299 CLPSMNDPQFGSCETIGFKENSTDYLYPEOLKMTYVKLISHRCQOPHYGSEVTTK 358
 Db 301 CLPSMNDPQFGSCETIGFKENSTDYLYPEOLKMTYVKLISHRCQOPHYGSEVTTK 360
 QY 359 MLCADPQMTDSCQDSDGSGPLVCSLOGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPW 418
 Db 361 MLCADPQMTDSCQDSDGSGPLVCSLOGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPW 420
 QY 419 IRSHTEKENGIAL 431
 Db 421 INTHTRGELNVL 433

RESULT 5

u-plasminogen activator (EC 3.4.21.73) precursor - rat
 N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text_change 18-Jun-1999
 C:Accession: S24604; 160186; 153472; S18932
 R:Rabdan, S.A.
 Submitted to the EMBL Data Library, April 1992
 A:Reference number: S24604
 A:Accession: S24604
 A:Molecule type: mRNA
 A:Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
 A:Cross-references: EMBL:X65651; NID:957456; PIDN:CAA4601.1; PID:957457
 A:Experimental source: tissue kidney
 R:Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
 Cancer Res. 52, 2489-2496, 1992
 A:Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
 A:Reference number: 160186; MUID:92233409; PMID:1568219
 A:Accession: 160186
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-432 <RES>
 A:Cross-references: EMBL:X63434; NID:957465; PIDN:CAA45028.1; PID:957466
 A:Experimental source: strain Fischer 344; tissue mammary
 R:Regno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blas, F.; Rossi, G.
 FEBS Lett. 306, 193-198, 1992
 A:Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
 A:Reference number: 153472; MUID:92339549; PMID:1321734
 A:Accession: 153472
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 31-62 <RES2>
 A:Cross-references: EMBL:X66907; NID:9396200; PIDN:CAA47356.1; PID:9396279
 C:Genetics:
 A:Gene: uPA
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:31-62/Domain: EGF homology <EGF>
 F:70-151/Domain: kringle homology <KR>
 F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

F:179-420/Domain: trypsin homology <TRY>
 F:168-300, 210-226, 218-289, 314-383, 346-362, 373-401/Disulfide bonds: #status predicted
 F:225, 276, 377/Active site: His, Asp, Ser #status predicted

Query Match 72.2%; Score 1728.5; DB 1; Length 432;
 Best Local Similarity 71.3%; Pred. No. 1,1e-124;
 Matches 308; Conservative 51; Mismatches 70; Indels 3; Gaps 3;

QY 1 MRALLARLLICVAVSDSKSGSNEIH-QVPENCCLNGTGVSKRYFSNIMHCNCPKFKG 59
 Db 1 MRVWLASFLICAL-VANSEGGSELEASDESNCCQNGVAVSYKYSIRCCCPKFKG 59
 QY 60 QHCEIDSKTCYGNHGFYRGKASTDTMGRPCLPMNSATVLOQTTHAHRSDALQLGK 119
 Db 60 EHCEIDSKTCYGNHGFYRGKASTDTMGRPCLPMNSATVLOQTTHAHRSDALQLGK 119
 QY 120 HNYCRNPDRRPMPCYVQVGLKPLVOCMHDCAQDKKPSPPPELKFQCGQKTLRPRFK 179
 Db 120 HNYCRNPDRRPMPCYVQVGLKPLVOCMHDCAQDKKPSPPPELKFQCGQKTLRPRFK 179
 QY 180 IIGEFITIEQPMFAIYRRHGGSYVYCGGSLISPCWVISAATHGIDYPKKEDYIVY 238
 Db 180 VGEFITYENQPMFAIYRRHGGSYVYCGGSLISPCWVISAATHGIDYPKKEDYIVY 239
 QY 239 LGRSLNSNTQGEKFEVENLILHKDYSADTLAHNDIALIKIRSKRGCAQPSRTIQT 298
 Db 240 LGQSLNSNTQGEKFEVENLILHKDYSADTLAHNDIALIKIRSKRGCAQPSRTIQT 299
 QY 299 CLPSMNDPQFGSCETIGFKENSTDYLYPEOLKMTYVKLISHRCQOPHYGSEVTTK 358
 Db 300 CLPSMNDPQFGSCETIGFKENSTDYLYPEOLKMTYVKLISHRCQOPHYGSEVTTK 359
 QY 359 MLCADPQMTDSCQDSDGSGPLVCSLOGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPW 418
 Db 360 MLCADPQMTDSCQDSDGSGPLVCSLOGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPW 419
 QY 419 IRSHTEKENGIAL 430
 Db 420 IQSHIGENGIAL 431

RESULT 6

u-plasminogen activator (EC 3.4.21.73) precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1987 #sequence revision 30-Sep-1987 #text_change 18-Jun-1999
 C:Accession: A29420; A24615
 R:Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
 Biochemistry 26, 8270-8279, 1987
 A:Title: The murine urokinase-type plasminogen activator gene.
 A:Reference number: A29420; MUID:88163489; PMID:2831940
 A:Accession: A29420
 A:Molecule type: DNA
 A:Residues: 1-433 <DBG>
 A:Cross-references: GB:J17922; NID:9202296; PIDN:AAA0539.1; PID:9202297
 R:Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher,
 Eur. J. Biochem. 148, 225-232, 1985
 A:Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase-
 A:Reference number: A24615; MUID:85179474; PMID:2885383
 A:Accession: A24615
 A:Molecule type: mRNA
 A:Residues: 1-433 <BEL>
 A:Cross-references: GB:X02389; NID:955127; PIDN:CAA26231.1; PID:955128
 C:Genetics:
 A:Insertions: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:32-63/Domain: EGF homology <EGF>
 F:71-152/Domain: kringle homology <KR>
 F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:180-421/Domain: trypsin homology <TRY>

F:169-301,211-227,219-290,315-384,347-363,374-402/disulfide bonds: #status predicted
F:226,277,378/Active site: His, Asp, Ser #status predicted

71.0%; Score 1700; DB 1; Length 433;
Best Local Similarity 69.2%; Pred. No. 1,76-122;

Matches 299; Conservative 56; Mismatches 75; Indels 2; Gaps 2;

QY 1 MRALLARLLCVLVSDSKSGSNELHGV-PSNCDLNGTCVSNKFSNIHMCNCPKRG 59
DB 1 MKWIALFLCALVVKNSSEGSVLGAPDESNGCGONGGVCSYKXFSIRSCSPRKEG 60
QY 60 OHGIDSKTCYBENGHFYRGKASTDPMGRCLPMNSATVLCQTYHAHRSALDGLGKH 119
DB 61 EHCIDASKTCYHNGSGSYRGKANTDTGRCLAMNAPVLYQKPNMHRPAISLGIGKH 120
QY 120 NYSNPNRRRRPWCYVGVGLKPLVQECVHDCADKKPSPPEELKFOCGQKTLRPFKI 179
DB 121 NYSNPNRRRRPWCYVGVGLKPLVQECVHDCADKKPSPPEELKFOCGQKTLRPFKI 180
QY 180 IGSEFTTIENQWPAIYRRHGGV-VTVYCGSLISPCWVISAHCFIDYPKKEDYIV 238
DB 181 VGGSEFTVENQWPAIYQKNGKSSPPSKCGSLISPCWVISAHCFIDYPKKEDYIV 240
QY 239 LGSRLNSNTQGEKFEVENLILHKDYADTLAHNDIALKIRSKGRCAPSRITQI 298
DB 241 LGSRLNSNTQGEKFEVENLILHKDYADTLAHNDIALKIRSKGRCAPSRITQI 300
QY 299 CLPSMYNDPQGTSCETIGFKENSTDYLYPEQLKMTVVLISHRECOQPHYGSEVT 358
DB 301 CLPSMYNDPQGTSCETIGFKENSTDYLYPEQLKMTVVLISHRECOQPHYGSEVT 360
QY 359 MLCADPQWKTDSGCGSGPLVCSLQGRMTLTGIVSGRGALDKRQVYTRVSHPLW 418
DB 361 MLCADPQWKTDSGCGSGPLVCSLQGRMTLTGIVSGRGALDKRQVYTRVSHPLW 420
QY 419 ITRHTKEENGLA 430
DB 421 IQSHIGKEGIA 432

RESULT 7

A35005
u-plasminogen activator (EC 3.4.21.73) precursor - chicken
N/Alternate names: uPA
C/Species: Gallus gallus (chicken)
C/Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C/Accession: A35005
R/Residue: N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A/Title: The chicken urokinase-type plasminogen activator gene.
A/Reference number: A35005; MUID:90110185; PMID:2295632
A/Accession: A35005
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-434 <ES>
A/Cross-references: GB:005187; NID:9212858; PIDN:AAA49131.1; PID:9212859
C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-111/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:40-71/Domain: EGF homology <EGF>
F:79-158/Domain: kringle homology <KR>
F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:173-416/Domain: trypsin homology <TRY>
F:162-296,202-218,210-285,310-379,342-350,369-397/Disulfide bonds: #status predicted
F:212,272,373/Active site: His, Asp, Ser #status predicted

Query Match 42.1%; Score 1009; DB 1; Length 434;
Best Local Similarity 46.0%; Pred. No. 1,2e-69;

Matches 196; Conservative 66; Mismatches 134; Indels 30; Gaps 8;

QY 10 LCVLVVS-DS-----KGSNEHQVPSNCDLNGTCVSNKFSNIHMCNCPKRG 61
DB 11 LCVLVVS-DS-----KGSNEHQVPSNCDLNGTCVSNKFSNIHMCNCPKRG 61

DB 11 LCVLVGSDVYIRQYKLSHKHRRPDRRECOCLNGTCITTRFSSQIKRCLCPBRYGGLH 70
QY 62 CEIDSKTCYBENGHFYRGKASTDPMGRCLPMNSATVLCQTYHAHRSALDGLGKH 120
DB 71 CEIDSKTCYBENGHFYRGKASTDPMGRCLPMNSATVLCQTYHAHRSALDGLGKH 126
QY 121 YCRNPNRRRRPWCYVGVGLKPLVQECVHDCADKKPSPPEELKFOCGQKTLRPFKI 180
DB 127 YCRNPNRRRRPWCYVGVGLKPLVQECVHDCADKKPSPPEELKFOCGQKTLRPFKI 174
QY 181 GSEFTTIENQWPAIYRRHGGV-VTVYCGSLISPCWVISAHCFIDYPKKEDYIV 236
DB 175 GSGQAEVETPWTAGIFQNTM-ITDQPLCGSLIDPCWVISAHCFIDYPKKEDYIV 233
QY 237 VYLSRLNSNTQGEKFEVENLILHKDYADTLAHNDIALKIRSKGRCAPSRITQI 296
DB 234 VYLSRLNSNTQGEKFEVENLILHKDYADTLAHNDIALKIRSKGRCAPSRITQI 293
QY 297 TICPSMYNDPQGTSCETIGFKENSTDYLYPEQLKMTVVLISHRECOQPHYGSEVT 356
DB 294 TICPSMYNDPQGTSCETIGFKENSTDYLYPEQLKMTVVLISHRECOQPHYGSEVT 353
QY 357 TKNLCADPQWKTDSGCGSGPLVCSLQGRMTLTGIVSGRGALDKRQVYTRVSHPL 416
DB 354 DNWVCADPQWKTDSGCGSGPLVCSLQGRMTLTGIVSGRGALDKRQVYTRVSHPL 413
QY 417 PMIRSH 422
DB 414 NMIDSN 419

RESULT 8

JS0598
t-plasminogen activator (EC 3.4.21.66) alpha-2 precursor - common vampire bat
N/Alternate names: tissue plasminogen activator
C/Species: Desmodus rotundus (common vampire bat)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C/Accession: JS0598
R/Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Dom
Gene 105, 229-237, 1991
A/Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A/Reference number: JS0597; MUID:92039036; PMID:1937019
A/Accession: JS0598
A/Molecule type: mRNA
A/Residues: 1-477 <KRA>
A/Cross-references: GB:M63988; NID:9166074; PIDN:AAA1593.1; PID:9166075
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F:47-79/Domain: fibronectin type I repeat homology <IRA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KR>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:185,398/Binding site: carbohydrate (Agn) (covalent) #status predicted
F:225-326/Cleavage site: His-Ser (plasmin) #status predicted
F:212,321,428/Active site: His, Asp, Ser #status predicted

Query Match 36.3%; Score 868.5; DB 2; Length 477;
Best Local Similarity 42.7%; Pred. No. 7.6e-59;
Matches 179; Conservative 61; Mismatches 150; Indels 29; Gaps 9;

QY 19 KGSNEHQVPSNCDLNGTCVSNKFSNIHMCNCPKRG 74
DB 74 RGLAQCHTVPVKSCSELRNCGTCQWASFSDF-VQCGKGYTGKCEVDTHATCYKDQ 132
QY 75 GHYRGASTDPMGRCLPMNSATVLCQTYHAHRSALDGLGKH YCRNPNRRRRPWCY 134
DB 133 GVTYRGWTSSESQAQCIWNSMLTRRTYNGRSDIITGLGNHNYCRNPNRRRRPWCY 132
QY 135 VQVGLKPLVQECVHDCADKKPSPPEELKFOCG-QKTLRPFKIIGSEFTTIENQW 193

Db 193 VKASFFILEFCSVPVCS-----KATCGLRKKEPQLHSTGGLFTDITSHPMQ 240
 194 AAIYRHRGGS-VTVYCGGSLISPCWVJISATGCFID-YKKEEDYIVYLGRSLNNTQGE 251
 Db 241 AAIYFONRNRSSSERFLCGSILSSCVLTPAHCFFERRPPQLRLRV-LGRYRVKGRKE 299
 252 MKFEVENLTIHKDYSADTLAHNDIALTLKIRSEKRCAPSRRTIOTICLPMYNDPQGT 311
 Db 300 GTFEVEKCI VHEPEDDT--YNNDIALLQKSGSPQCGESDVSVAICLPKANLQLPWT 357
 312 SEITGFGKENVSTDIYVLPQQLMTYVVKLISHRECCQPHYYSSEVTTKMLCAADPQWKT-- 369
 Db 358 ECEISGCGYGRKSSSPFYSQLEKGVHRYLPSSRCSKFLFNKTIVNNMLCAGDTRSGEILY 417
 370 ---DSGCGDSGGPVPVCSLQGRMTLTGIVSMGRGALXDKPQYRVHFLPMIRSHRK 424
 Db 418 PNVHACQCGDSGGLVLCNMNDNMTLLGLISWVGCGEKDIPGVYTKVTNVLGWIIRDNR 476

RESULT 9
 t-plasminogen activator (BC 3.4.21.68) precursor [validated] - human
 N:Alternate names: t-PA, tissue plasminogen activator
 C:Species: Homo sapiens (man)
 C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Dec-2000
 C:Accession: A94004; A23529; J70562; A93293; S02125; A91343; A93951; A91322; A54645; I60
 R:NY, T., Eligh, F., Lund, B.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
 A:Title: The structure of the human tissue-type plasminogen activator gene: correlation
 A:Reference number: A94004; MUID:84298137; PMID:6089198
 A:Accession: A94004
 A:Molecule type: DNA
 A:Residues: 1-562 <N>
 A:Cross-references: GB:I00141
 A:Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translac
 R:Frietzer Degen, S.J.; Rajput, B.; Reich, E.
 J. Biol. Chem. 261, 6972-6985, 1986
 A:Title: The human tissue plasminogen activator gene.
 A:Reference number: A23529; MUID:86196143; PMID:3009482
 A:Accession: A23529
 A:Molecule type: DNA
 A:Residues: 1-562 <D>
 A:Cross-references: GB:K03021; NID:g339817; PIDN:AAA98809.1; PID:g339818
 R:Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.
 Agric. Biol. Chem. 55, 1225-1232, 1991
 A:Title: Purification and characterization of tissue plasminogen activator secreted by h
 A:Reference number: J70562; MUID:91291340; PMID:1368681
 A:Accession: J70562
 A:Molecule type: mRNA
 A:Residues: 31-562 <T>
 A:Cross-references: DDBJ:D01096; NID:g220128; PIDN:BAA00881.1; PID:g441174
 A:Experimental source: embryonic lung fibroblast IMR-90 cells
 A:Note: part of this sequence, including the amino end of the mature protein, was confir
 R:Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harting, R.N.; Vohar, G.A.; Ward, C.A.; Bennett
 Nature 301, 214-221, 1983
 A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
 A:Reference number: A93293; MUID:83115262; PMID:6337343
 A:Accession: A93293
 A:Molecule type: mRNA
 A:Residues: 1-562 <P>
 A:Cross-references: GB:L00141
 A:Experimental source: melanoma cells
 R:Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
 Nucleic Acids Res. 16, 5695, 1988
 A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
 A:Reference number: S02125; MUID:88265579; PMID:3133640
 A:Accession: S02125
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-562 <S>
 A:Cross-references: EMBL:X07393; NID:g37243; PIDN:CMA30302.1; PID:g37244
 A:Experimental source: fetal lung cells

R:Kagiani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Mat
 FEBS Lett. 189, 145-149, 1985
 A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
 A:Reference number: A91343; MUID:85285620; PMID:3896853
 A:Accession: A91343
 A:Molecule type: mRNA
 A:Residues: 1-38, 'G', 86-433, 'E', 435-562 <K>
 A:Experimental source: Detroit 562 cells; ATCC 138
 R:Edlund, T.; Ny, T.; Rando, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
 Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
 A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
 A:Reference number: A93951; MUID:83169656; PMID:6572897
 A:Accession: A93951
 A:Molecule type: mRNA
 A:Residues: 251-358 <E>
 A:Experimental source: melanoma cells
 R:Poll, G.; Kallstrom, M.; Bergsdoerf, N.; Wallen, P.; Jornvall, H.
 Biochemistry 23, 3701-3707, 1984
 A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
 differences.
 A:Reference number: A90488; MUID:85000469; PMID:6433976
 A:Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and i
 R:Poll, G.; Kaplan, L.; Elhansson, M.; Wallen, P.; Jornvall, H.
 FEBS Lett. 168, 29-32, 1984
 A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
 A:Reference number: A91322; MUID:84158956; PMID:6538514
 A:Accession: A91322
 A:Molecule type: protein
 A:Residues: 33-45/311-320 <P>
 A:Experimental source: uterus
 A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
 R:van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
 J. Biol. Chem. 261, 14214-14218, 1986
 A:Reference number: A37567; MUID:87033611; PMID:3021732
 A:Contents: annotation; fibrin binding site
 R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
 EMBO J. 5, 3525-3530, 1986
 A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen act
 A:Reference number: A37568; MUID:87161761; PMID:3030730
 A:Contents: annotation; fibrin binding site
 R:Dodd, I.; Nunn, B.; Robinson, J.H.
 Thromb. Haemost. 59, 523-528, 1988
 A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type p
 A:Reference number: A60902; MUID:89044681; PMID:3142086
 A:Contents: annotation; novel forms of expressed recombinant t-PA
 R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenacker, G.; V
 Mol. Biol. Med. 3, 279-292, 1986
 A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expre
 A:Reference number: A54645; MUID:86284200; PMID:3090401
 A:Accession: A54645
 A:Molecule type: mRNA
 A:Residues: 1-562 <H>
 A:Cross-references: GB:M1518; NID:g190031; PIDN:AAA6011.1; PID:g190032
 A:Note: parts of this sequence were confirmed by peptide sequencing
 R:Reddy, V.B.; Garrimone, A.J.; Sasaki, H.; Wei, C.
 DNA 6, 461-472, 1987
 A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
 A:Reference number: I60110; MUID:88054470; PMID:2824147
 A:Accession: I60110
 A:Molecule type: translated from GB/EMBL/DBJ
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-562 <R>
 A:Cross-references: GB:M18182; NID:g340176; PIDN:AAA6800.1; PID:g340177
 R:Fishar, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schlenning, W.D.
 J. Biol. Chem. 260, 11223-11230, 1985
 A:Title: Isolation and characterization of the human tissue-type plasminogen activator st
 A:Reference number: I55232; MUID:85289338; PMID:3161893
 A:Accession: I55232
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-36 <R>
 A:Cross-references: GB:M1890; NID:g339837; PIDN:AAA61213.1; PID:g339839
 C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single

Query Match	36.3%;	Score 868.5;	DB 1;	Length 562;
Best Local Similarity	37.9%;	Pred. No. 9.1e-59;		
Matches 187;	Conservative 57;	Mismatches 160;	Indels 89;	Gaps 11

RESULT 10

t-plasminogen activator (EC 3.4.21.66) beta precursor - common vampire bat
N/Alternate names: tissue plasminogen activator
C/Species: *Desmodus rotundus* (common vampire bat)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #next_change 16-Jul-1999

Query Match	36.2%;	Score 867.5;	DB 2;	Length 431;
Best Local Similarity	41.6%;	Pred. No. 8,1e-59;		
Matches 185;	Conservative 61;	Mismatches 152;	Indels 47;	Gaps 10

RESULT 11

t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C.Species: Megaderma lyra
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #ext_change 10-Sep-1999
C.Accession: A34369
R.Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacobs
J. Biol. Chem. 264, 19947-17952, 1989
A.Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasminic
A.Reference number: A34369; MUID:90036867; PMID:2509450
A.Accession: A34369
A.Status: preliminary
A.Molecule type: mRNA

A:Residues: 1-477 <GAR>
A:Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KR>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 36.2%; Score 867.5; DB 1; Length 477;
Best Local Similarity 42.7%; Pred. No. 9,1e-59;
Matches 179; Conservative 61; Mismatches 150; Indels 29; Gaps 9;

QY 19 KGSNELHQVP-----SNCDCLNGTGVSNKYPFSNIMHCNCPKKFGQHCEIDSKTCYEGN 74
Db RGLAQCHTVPVKSCSELRFCNGTGWQASFSDF-VCCPRKGYTGKCEVDTHATCYKQ 132
QY 75 GHFPRGKASTDTMRPCLPANSAIVLQOTYHAHSDALQLGKHNKCRNDNRRCW 134
Db 133 GVTVRGWSTSESGAQCINMNSNLTTRTYNGRSDAITGLGNHNYCRNDNNSKWCY 192
QY 135 VQVGLKPLVQECWHDCAADGKPPPEBELKFCQG-OKTLPFRKTIIGSEPTIENQPF 193
Db 193 VIKASKTILERCSPVCS-----KATGRLKRYKRPQLHSGGLFTDITSHPW 240
QY 194 AAIYRRHGGG-VTVVCGSLISPCWYISATHCFID-YPKKEDYIVYLGRSLNSNTQGE 251
Db 241 AAIYRRHGGG-VTVVCGSLISPCWYISATHCFID-YPKKEDYIVYLGRSLNSNTQGE 251
QY 252 MKFEVENLILHKYASDTLHNDILALIKRSKGRCAQSPRTQITCLPMYNDPOFCT 311
Db 300 QTFVEVEKCTVHEEEDDT--YNDIALQLKSGSPQCAQESDSVRAICLPANQLPDMT 357
QY 312 SCEITFGKENSSTYLYPEOLKMTVVKLISHRECOQPHYVSEVTTKMLCAADPOMKT-- 369
Db 358 ECEISGTYGKSSPFPISSEQLKEGHVRLYPSRCTSKPLFKIVTYTKMLCGDTRSGEIH 417
QY 370 ----DSCQDGSGBPLVCSLQGRMTLTGIVSWGRGKALDKDPVYTRYVHFLPWIRSHTK 424
Db 418 PNVHADACQDGSGBPLVCSLQGRMTLTGIVSWGRGKALDKDPVYTRYVHFLPWIRSHTK 476

RESULT 12

A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35029; A31597
R:Feing, P.; Ohlsson, M.; NY, T.
J. Biol. Chem. 265, 2022-2027, 1990
A:Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec
A:Reference number: A35029; MUID:90130448; PMID:2105315
A:Accession: A35029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-559 <FEN>
A:Cross-references: GB:M31197; NID:g207429; PIDN:AAA42261.1; PID:g207431; GB:J05226
R:NY, T.; Leonardson, G.; Hsueh, A.J.W.
DNA 7, 671-677, 1988
A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activat
A:Reference number: A31597; MUID:89170114; PMID:3148445
A:Accession: A31597
A:Molecule type: mRNA
A:Residues: 1-379, 'K', 381-559 <NYT>
A:Cross-references: GB:M23667; NID:g530159; PIDN:AAA4812.1; PID:g530160
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <1FI>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-394/Domain: kringle homology <KR2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F:149,481/Binding site: carbohydrate (Aan) (covalent) #status predicted
F:308-309/Cleavage site: Arg-ile (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 35.4%; Score 847.5; DB 1; Length 559;
Best Local Similarity 36.7%; Pred. No. 3,6e-57;
Matches 183; Conservative 65; Mismatches 153; Indels 97; Gaps 12;

QY 18 SKGSNELHQVP-----SNCDCLNGTGVSNKYPFSNIMHCNCPKKFGQHCEIDSKTCYEG 73
Db 69 NSGLAQCHSVPVKSCSEPRFCNGTGWQALYFSDF-VCCPRDFVGRKCDIDTRATCFEG 127
QY 74 NGHPRGKASTDTMRPCLPANSAIVLQOTYHAHSDALQLGKHNKCRNDNRRCW 133
Db 128 QGTVRGWSTSESGAQCINMNSNLTTRTYNGRSDAITGLGNHNYCRNDNNSKWCY 187
QY 134 VQVGLKPLVQECWHDCAADGKPPPEBELKFCQG-VHDCADGK-----KPSPP----- 161
Db 188 YVFKAKYITTEFCSTACPKFPEDCYGKGVTRYRGHSTTSKASCLPMNSMILGKTY 247
QY 162 -----BEIKRQ-----CGQKTLR-PRF 177
Db 248 TAWRANSQALGLGRHNYCRNDNADGKPPPEBELKFCQG-VHDCADGK-----KPSPP----- 161
QY 178 KIIGSEPTIENQPFPAIY-RHRRGASVTVVCGSLISPCWYISATHCFIDYPKKEDYI 236
Db 308 RIRGGFTDITSHPWQALFVKNRSPGERFLCGVLISSCWLSAHCHEVERPPPHHLK 367
QY 237 VVGRSLNSNTQGEKMEFEVENLILHKYASDTLHNDILALIKRSKGRCAQSPRTQ 296
Db 368 VVGRSLNSNTQGEKMEFEVENLILHKYASDTLHNDILALIKRSKGRCAQSPRTQ 296
QY 297 TICLPSTNDPOF-----GTSCEITFGKENSSTYLYPEOLKMTVVKLISHRECOQPHYV 352
Db 426 TACLP-----DPVQLPDMTECELSGKHEASPPFSRDLKHAHVLVPSRCTSGHLEFN 481
QY 353 SEVTTKMLCAADP-----QKTDSCQDGSGBPLVCSLQGRMTLTGIVSWGRGKALDKDPG 407
Db 482 KTITSNMLCGDTRTGNDVHACQDGSGBPLVCMIDKRMILLGLISWGLGCGQKDPG 541
QY 408 VYTRVSHFLPWIRSHTK 425
Db 542 IYTKVTVYLVNWDNMKQ 559

RESULT 13

JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0597
R:Krauszschmar, J.; Haendler, B.; Langer, G.; Bojdl, W.; Bringmann, P.; Alagon, A.; Dom
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0597
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63987; NID:g166070; PIDN:AAA31591.1; PID:g166071
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>

F/22-36/Domain: propeptide #status predicted <PRO>
 F/37-47/Product: plasminogen activator alpha-1 #status predicted <PLA>
 F/42-79/Domain: fibronectin type I repeat homology <1FA>
 F/87-120/Domain: EGF homology <EGF>
 F/128-209/Domain: kring homology <KR>
 F/226-471/Domain: trypsin homology <TRY>
 F/42-79,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-245,257-273,265-334,359-4
 F/153,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F/225-226/Cleavage site: His-Ser (plasmin) #status predicted
 F/272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 35.2%; Score 841.5; DB 2; Length 477;
 Best Local Similarity 41.8%; Pred. No. 8.8e-57;
 Matches 176; Conservative 57; Mismatches 151; Indels 37; Gaps 10;

19 KGSNELHQPVSN-CD---CLNGGTCVSNKYPFSNHNWNCNPKKGGCHGCHIDSKTCYEG 74
 74 KQARCHTVVNSCSCEPCNCGTCWQAVYFSDP-VCCPCGPGYTGKCEVDTRATCYEQ 132
 75 GHFPRKASTDTMGRPCLPMNSATVLCQTYAHRSALQGLGKHNYCRNPNRRPWCY 134
 133 GVTYRGTWSTABSRVVEGINNSSLTRTYNGRMDAFNIGLGNHYCRPNNGAPKPCWCY 192
 135 VQVGLKPLVQECMVHDCADGKKSSPPEELKFGCG-OXTLRPRFKIIGSEFTTINQPF 193
 193 VIKAGKFTSESCSPVCS-----KATGRLKRYKEPOLHSTGGLFTDITSHPMQ 240
 194 AAIYRRHRRGS-VTYVGGSLISPCWYISATHGCIIDPKKEDY-----VILGRSLNSN 247
 241 AAIYRRHRRGS-VTYVGGSLISPCWYISATHGCIIDPKKEDY-----VILGRSLNSN 247
 248 TQSEMKFEVENLILHKDYSADTLAHNDIALKIRSKRCAQPSRTIQTICLPSMYNDP 307
 296 GEEGQFQKVKYIVHKEFDDT--YNNDIALLQKSSPCQAGSDSVRAICLPEALQL 353
 308 QGFTSCETGFGKENSVDLYLPOLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQW 367
 354 PDWTECELSGKSSPFYSQOLKEGHVRLYPSRCAPKFLPNKVTNNMLCAGDTRS 413
 368 KT-----DSGCGDSGGLVCSLQGMRTLGIYSWKGKCALKDKPGYTVSHFLPIRS 421
 414 GEIYPMVHDCQSDSGGLVCMNDHMTLLGIISWVGCEKQDPGVYTVTVYLGWIRD 473
 422 H 422
 474 N 474

RESULT 14

A29941
 t-plasminogen activator (EC 3.4.21.68) precursor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: A29941; S48205; S48206
 R/Rickles, R.J.; Darrow, A.L.; Strickland, S.
 J. Biol. Chem. 263, 1563-1569, 1988
 A/Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator m
 A/Reference number: A29941; MUID:88087303; PMID:2826484
 A/Accession: A29941
 A/Molecule type: mRNA
 A/Residues: 1-559 <RIC>
 A/Cross-references: GB:J03520; NID:G202109; PIDN:AAA40470.1; PID:G202110
 R/Liljen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
 Eur. J. Biochem. 224, 863-871, 1994
 A/Title: Characterization of the murine plasma fibrinolytic system.
 A/Reference number: S48202; MUID:95010076; PMID:7523120
 A/Accession: S48205
 A/Molecule type: protein
 A/Residues: 33-37,'X',39-40 <LIJ>
 A/Accession: S48207
 A/Molecule type: protein
 A/Residues: 309-316 <LIZ>
 A/Accession: S48206

A/Molecule type: protein
 A/Residues: 33-37,'X',39-40 <LIW>
 C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 A/Keywords: fibrinolysis; glycoprotein; hydrolase; kring; serine proteinase
 F/1-17/Domain: signal sequence #status predicted <SIG>
 F/18-29/Domain: propeptide #status predicted <PRO>
 F/30-559/Product: t-plasminogen activator #status predicted <MAT>
 F/30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
 F/149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F/308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
 F/355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 34.8%; Score 833.5; DB 1; Length 559;
 Best Local Similarity 35.7%; Pred. No. 4.3e-56;
 Matches 178; Conservative 66; Mismatches 157; Indels 97; Gaps 10;

18 SKGSNELHQPVSN-CD---CLNGGTCVSNKYPFSNHNWNCNPKKGGCHGCHIDSKTCYEG 73
 69 NSGLVQCHSVPRVSCSEPCNCGTCQALYFSDP-VCCPCGPGYTGKCEVDTRATCYEQ 127
 74 NGHPRKASTDTMGRPCLPMNSATVLCQTYAHRSALQGLGKHNYCRNPNRRPWCY 133
 128 QGTYRGTWSTABSRVVEGINNSSLTRTYNGRMDAFNIGLGNHYCRPNNGAPKPCWC 187
 134 VQVGLKPLVQECMVHDCADGKKSSPPEELKFGCG-OXTLRPRFKIIGSEFTTINQPF 193
 188 YVRKAKYTFEPCSTACPRKSGEDCVGKGYTRGTHSLTTSQASCLPMNSIYLMGSKY 247
 159 -----SPPEELK-----FQCG-OXTLRPRF 177
 248 TAWRTNSQALGLARHNYCRNPDGDARPCVMDKRLTWECDSMPCSTGLRQYKRPQF 307
 178 KIIGSEFTTINQPMFPAIY-RHRRGSVTVVGGSLISPCWYISATHGCIIDPKKEDY 236
 308 RINGSLYTDITSHPMQALFVKNKRSPGERELCGVLISSCWLSAACHLELPPNNHLK 367
 237 VILGRSLNSNTQSEMKFEVENLILHKDYSADTLAHNDIALKIRSKRCAQPSRTIQ 236
 368 VILGRYRVVPGSEEGQFLEKTYVHEFDDT--YNDIALLQKSSPCQAGSDSVRAICLPEALQL 425
 297 TILPSMYNDPQF-----GTSCETGFGKENSVDLYLPOLKMTVVKLISHRECOQPHYG 352
 426 TACLP-----DPNQLPDWTECELSGKSSPFSDRLKEAHVRLYPSRCTSQHLFN 481
 353 SEVTTKMLCAADP-----QMKTDSGCGDSGGLVCSLQGMRTLGIYSWKGKCALKDKPG 407
 482 KYTVNNMLCAGDTRSGNDLHDCQSDSGGLVCMINQMUTLGIISWGLGCGQKDPVG 541
 408 VYTVRSHFLPMIRSHTKK 425
 542 VYTKVTNYIDMIDNMKQ 559

RESULT 15

US0600
 t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
 N/Alternate names: tissue plasminogen activator
 C/Species: Desmodus rotundus (common vampire bat)
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 A/Accession: US0600
 R/Kraetzschmar, J.; Haendler, B.; Langer, G.; Bofdoll, W.; Bringmann, P.; Alagon, A.; Donr
 Gene 105, 229-237, 1991
 A/Title: The plasminogen activator family from the salivary gland of the vampire bat Desm
 A/Reference number: US0597; MUID:92039036; PMID:1937019
 A/Accession: US0600
 A/Molecule type: mRNA

A:Residues: 1-394 <KRA>
A:Cross-references: GB:M63990; NID:G16078; PIDN:AAA1595.1; PID:G16079
A:Note: the authors translated the codon ATC for residue 75 as Thr
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:122-36/Domain: propeptide #status predicted <PRO>
F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
F:143-388/Domain: kringle homology <KR>
F:45-126/66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
F:189,238,345/Active site: His, Asp, Ser #status predicted
F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.4%; Score 752; DB 2; Length 394;
Best Local Similarity 42.0%; Pred. No. 5,1e-50;
Matches 155; Conservative 56; Mismatches 134; Indels 24; Gaps 7;
QY 65 DKSRTCYEGNGHFYRGKASTDTMGSPCLPWN SATVLQOTYHAHRS DALQLGIGHNYCRN 124
Db 40 DPHATCYDQGVTRGTWTSTESGAQCTINWNSLIRITVGRMPEAVKLGHNHYCRN 99
QY 125 PDNRPRPCTYQVGLKPLVQECWVHDCADGKKPSSPPEELKFCG-QKTRPRFKITIGE 183
Db 100 PDGASKPWCYVTKARKFTSESCSVPCS-----KATGRLRYKEPQLHSTGGL 147
QY 184 FTTIENOPFAIYRRHRS -VTYVCGSLISPCWVISA THCFID- YPKKEDIYVILGR 241
Db 148 FTDITSHPMQALFAQNRRSSGERFLCGILISSCWVLTAAHCFQERYPPQHLRV-LGR 206
QY 242 SRLNSNTQGEKFEVENILHKDYADTLAHNDIALIKIRSKERCAQPSRTIQTCLP 301
Db 207 TYRVKPGKEQTFVEKCTVHEEFDPT- YNNDIALQLKSGSPQCAQESDSVRAICLP 264
QY 302 SMYNDPQFGTCEITGPKENSTDYLPQULMTVVKLISHRCQOPHYGSEVTTMLC 361
Db 265 EANLQLPDMTECELSGYGKHS SPFSYEQLEKGVRLYPSSRCTSKFLFNKVTYNNMLC 324
QY 362 AADPOMKT-----DSCQDSGSPLVCSLQGRMTLTGIVSGRGCALKDXGVYTRVSHF 415
Db 325 AGDTRSGEIIYPNVDAICQDSGGFLVCNNDNMHTLLGIISWVGCGEKDIPGVYTKVINY 384
QY 416 LPWIRSHTK 424
Db 385 LGWIRDNMNR 393

Search completed: March 18, 2004, 11:20:16
Job time : 45.326 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:54:45 ; Search time 26.407 Seconds

(without alignments)
849.859 Million cell updates/sec

Title: US-10-076-421-2

Perfect score: 2394

Sequence: 1 MRALLRLLLCVLVSDSKG.....VSHFLPWRSHKENGALL 431

Scoring table: BLOSUM62

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2391	99.9	431	1 UROK_HUMAN	P00749 homo sapien
2	2228	93.1	433	1 UROK_PAPCY	P16227 papio cynoc
3	1928.5	80.6	442	1 UROK_PIG	P04185 sus scrofa
4	1835	76.6	433	1 UROK_BOVIN	Q05589 bos taurus
5	1728.5	72.2	432	1 UROK_RAT	P29598 rattus norv
6	1700	71.0	433	1 UROK_MOUSE	P06869 mus musculu
7	1009	42.1	434	1 UROK_CHICK	P15120 gallus gall
8	868.5	36.3	477	1 URT2_DESRO	P15638 desmodus ro
9	868.5	36.3	552	1 TPA_HUMAN	P00750 homo sapien
10	867.5	36.2	431	1 URTB_DESRO	P98121 desmodus ro
11	847.5	35.4	559	1 TPA_RAT	P19637 rattus norv
12	841.5	35.2	477	1 URT1_DESRO	P98119 desmodus ro
13	828.5	34.6	559	1 TPA_MOUSE	P11214 mus musculu
14	821	34.3	566	1 TPA_BOVIN	Q28198 bos taurus
15	752	31.4	394	1 URTG_DESRO	P49150 desmodus ro
16	735.5	30.7	654	1 HGFA_HUMAN	Q04756 homo sapien
17	727	30.4	653	1 HGFA_MOUSE	Q09088 mus musculu
18	718.5	30.0	603	1 FA12_CAVPO	Q04962 cavia porce
19	692	28.9	635	1 FA12_HUMAN	P00748 homo sapien
20	643.5	26.9	593	1 FA12_HUMAN	P98140 bos taurus
21	504.5	21.1	790	1 PLMN_PIG	P06867 sus scrofa
22	501	20.9	812	1 PLMN_MOUSE	P20918 mus musculu
23	499	20.8	810	1 PLMN_HUMAN	P00747 homo sapien
24	497	20.8	810	1 PLMN_MOUSE	P12545 macaca mula
25	494.5	20.7	333	1 PLMN_CANFA	P80009 canis fami
26	493	20.6	343	1 PLMN_SHEEP	P81286 ovis aries
27	484.5	20.2	812	1 PLMN_BOVIN	P06868 bos taurus
28	474	19.8	1420	1 APOA_MACMU	P14417 macaca mula
29	474	19.8	4548	1 APOA_HUMAN	P88519 homo sapien
30	462	19.3	875	1 NETR_HUMAN	P56730 homo sapien
31	449.5	18.8	338	1 PLMN_HORSE	P80010 equus cabal
32	445	18.6	761	1 NETR_MOUSE	O08762 mus musculu
33	435	18.2	810	1 PLMN_ERIEU	Q29485 erinaceus e

34	424.5	17.7	418	1 HAT1_HUMAN	O60235 homo sapien
35	420	17.5	811	1 TMS6_HUMAN	Q81800 mus sapien
36	419.5	17.5	811	1 TMS6_MOUSE	Q9db10 mus musculu
37	417	17.4	436	1 HEP5_MOUSE	O35453 mus musculu
38	410	17.1	638	1 KAL_MOUSE	P26262 mus musculu
39	408.5	17.1	343	1 PSS5_HUMAN	O16551 homo sapien
40	406	17.0	711	1 HGFL_HUMAN	P26927 homo sapien
41	405.5	16.9	855	1 ST14_HUMAN	Q9Y5Y6 homo sapien
42	405.5	16.9	855	1 ST14_MOUSE	P56677 mus musculu
43	405	16.9	277	1 KLKD_HUMAN	Q9UKR3 homo sapien
44	404.5	16.9	455	1 TMS5_MOUSE	Q9er04 mus musculu
45	402.5	16.8	417	1 HEP5_HUMAN	P05981 homo sapien

ALIGNMENTS

RESULT 1
ID UROK_HUMAN STANDARD; PRT; 431 AA.
AC P00749; Q15844; Q16618; Q969W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
GN (U-plasminogen activator).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=85215647; PubMed=2987867;
RX Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;
RA "The human urokinase-plasminogen activator gene and its promoter.";
RT Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RT Steffens G.J., Heyneker H.L.;
RA "Cloning and expression of the gene for pro-urokinase in Escherichia
coll.";
RT Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056954; PubMed=2415429;
RA Nagai W., Hirataken R., Kaneda T., Hayasuke N., Arimura H.,
RT "Molecular cloning of cDNA coding for human prepro-urokinase.";
RN Gene 36:183-188(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,
RT van Elsen A., Herzog A., Bollen A.;
RA "Molecular cloning, sequencing, and expression in Escherichia coli of
human prepro-urokinase cDNA.";
RN DNA 4:139-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RT Nickerson D.A.;
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT Stransberg R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsien F.,
RA Datchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carmini P., Prange C.,
 RA Bawa S.S., Loguano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Keltman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schermer A., Schein J.B., Jones S.J.M., Matra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [17]
 RP SEQUENCE OF 66-431 FROM N.A.
 RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Stoppeilli M.P., Galeffi P., di Nocera P., Biasi F.,
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA.";
 RT Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [18]
 RP SEQUENCE OF 21-177.
 RX MEDLINE=83055084; PubMed=6754569;
 RA Gunzler W.A., Steffens G.J., Oetting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.,
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RT Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [19]
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749491;
 RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,
 RA Studer R.O.,
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RT Eur. J. Biochem. 125:251-257(1982).
 RN [10]
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Oetting F., Frankus E., Flohe L.,
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RT Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1056(1982).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.,
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RT Structure 3:681-691(1995).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdolen V., Huber R., Moroder L.,
 RT "4-aminomethylphenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RT Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13]
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.,
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RT Nature 337:579-582(1989).
 RN [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.,

RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase.";
 RT Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=9419701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.,
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RT J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=9616279; PubMed=8652631;
 RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
 RA Sawaaki Y., Hanada K.,
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle
 RT structure.";
 RT Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Conne B., Berczy M., Belin D.,
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RT Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP ERRATUM.
 RA Conne B., Berczy M., Belin D.,
 RT Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmaloldt B., Trommler P., Hell W.,
 RA Creutzburg S., Graef H., Magdolen V.,
 RT "Molecular analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RT Electrophoresis 18:686-689(1997).
 RN [1]
 CC -1- FUNCTION: Potent plasminogen activator and is clinically used for
 CC therapy of thrombolytic disorders.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1 chain.
 CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
 CC in pulmonary embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 BGF-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X02419; CAA26268.1; -;
 DR EMBL: M15476; AAA61253.1; -;
 DR EMBL: D00244; BAA00175.1; -;
 DR EMBL: D11143; BAA01919.1; -;
 DR EMBL: X02760; CAA26535.1; -;
 DR EMBL: AF377330; AAK53822.1; -;
 DR EMBL: BC013575; AAH13575.1; -;
 DR EMBL: X03226; AAC97138.1; -;
 DR EMBL: K03286; AAA61252.1; -;
 DR EMBL: A21571; CAA01559.1; -;
 DR EMBL: A18397; CAA01390.1; -;
 DR PIR: A00931; UKHU.
 DR PDB: 1KDU; 31-OCT-93.

Query Match 99.4%; Score 2391; DB 1; Length 431;
 Best Local Similarity 99.8%; Pred. No. 6.6e-182;
 Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLLARLLLCVIVSDSKSNEHQVPSNCDLNGTCVSNKYFSNHNHNCPEKKGQ 60
 DB 1 MRLLARLLLCVIVSDSKSNEHQVPSNCDLNGTCVSNKYFSNHNHNCPEKKGQ 60
 QY HCEIDSKTCYBENGHFRKASTDTMGRPCLPWNSATVLAQYTHARSDALQLGSKN 120
 DB HCEIDSKTCYBENGHFRKASTDTMGRPCLPWNSATVLAQYTHARSDALQLGSKN 120
 QY 121 YCNPNRRRPMWCYVQGLKPLVQECVMDKADGKSPPEELKFCGQKTLRPRKII 180
 DB 121 YCNPNRRRPMWCYVQGLKPLVQECVMDKADGKSPPEELKFCGQKTLRPRKII 180
 QY 181 GGEFTTIENQPMFALYRRHGGSVTVYCGSLISPCWVISAHCIFIDYKEDYIYLG 240
 DB 181 GGEFTTIENQPMFALYRRHGGSVTVYCGSLISPCWVISAHCIFIDYKEDYIYLG 240
 QY 241 RSLNNTQSEMEFEVENLLHMDYADTLAHNDIALKIRSEGCAPSRPTOTICL 300
 DB 241 RSLNNTQSEMEFEVENLLHMDYADTLAHNDIALKIRSEGCAPSRPTOTICL 300
 QY 301 PSWYNDPOFGTSCETGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTKML 360
 DB 301 PSWYNDPOFGTSCETGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTKML 360
 QY 361 CAADPMKTDSCQDSSGGLVCSLQGMVLTGIVSWRGALADKDPGVTVRVSHFLWIR 420
 DB 361 CAADPMKTDSCQDSSGGLVCSLQGMVLTGIVSWRGALADKDPGVTVRVSHFLWIR 420
 QY 421 SHTKEENGLAL 431
 DB 421 SHTKEENGLAL 431

RESULT 2
 UROK_PAPCY STANDARD; PRT; 433 AA.

AC 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Papio cynocephalus (Yellow baboon).
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 NCBI_TaxID=9556;
 RN NCBI
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thoracic aorta;
 RX MEDLINE=90287734; PubMed=2113276;
 RA Au Y.P.T., Wang T.W., Clowes A.W.;
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
 RT plasminogen activator."
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1 chain (by similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).

CC -----
 CC EMBL; X51935; CAA6200.1; -.
 CC PIR; S14687; URBAY.
 CC HSSP; P00749; 1LMW.
 CC MEROPS; S01.231; -.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006209; EGF_like.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR008293; Pept_S1A_UPA.
 CC InterPro; IPR001254; Peptidase_S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC Pfam; PF00051; Kringle1.
 CC Pfam; PF00089; trypsin; 1.
 CC PIRSF; PIRSF001144; Urk_plasm_act; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 1.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; TYP_SPC; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; FALSE_NEG.
 CC PROSITE; PS00026; EGF_3; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS00240; KRINGLE_2; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC KX Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 CC KX Kringle; EGF-like domain; Zymogen; Signal.
 CC FT SIGNAL 1 20
 CC FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 CC FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
 CC FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
 CC FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
 CC FT CHAIN 26 62 EGF-LIKE.
 CC FT DOMAIN 69 150 KRINGLE.
 CC FT DOMAIN 151 177 CONNECTING PEPTIDE.
 CC FT DOMAIN 178 433 SERINE PROTEASE.
 CC FT DISULFID 30 38 BY SIMILARITY.
 CC FT DISULFID 32 50 BY SIMILARITY.
 CC FT DISULFID 52 61 BY SIMILARITY.
 CC FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
 CC FT DISULFID 208 224 BY SIMILARITY.
 CC FT DISULFID 216 287 BY SIMILARITY.
 CC FT DISULFID 315 384 BY SIMILARITY.
 CC FT DISULFID 347 402 BY SIMILARITY.
 CC FT ACT SITE 223 223 CHARGE RELAY SYSTEM.
 CC FT ACT SITE 224 274 CHARGE RELAY SYSTEM.
 CC FT ACT SITE 378 378 CHARGE RELAY SYSTEM.
 CC FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
 CC SQ SSQUNCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 93.1%; Score 2228; DB 1; Length 433;
 Best Local Similarity 92.4%; Pred. No. 5.2e-169;
 Matches 401; Conservative 17; Mismatches 12; Indels 4; Gaps 2;

QY 1 MRLLARLLLCVIVSDSKSNEHQVPSNCDLNGTCVSNKYFSNHNHNCPEKKGQ 60
 DB 1 MRLLARLLLCVIVSDSKSNEHQVPSNCDLNGTCVSNKYFSNHNHNCPEKKGQ 59
 QY HCEIDSKTCYBENGHFRKASTDTMGRPCLPWNSATVLAQYTHARSDALQLGSKN 120
 DB HCEIDSKTCYBENGHFRKASTDTMGRPCLPWNSATVLAQYTHARSDALQLGSKN 119
 QY 121 YCNPNRRRPMWCYVQGLKPLVQECVMDKADGKSPPEELKFCGQKTLRPRKII 180

```

Db 120 YCRNPDRRRPWCYVGLKQVQECMVHNCAGKKSPSEETQFCGQRTLRPRKIV 179
QY 181 GGEFTTLENQWFAIYRRRRGGSVTVVCGSLISPCWVLSATHCFIDYKKEDEYIYLG 240
Db 180 GGEFTTLENQWFAIYRRRRGGSVTVVCGSLISPCWVLSATHCFIDYKKEDEYIYLG 239
QY 241 RSRLLNSTQCEMKFEVENLILHKDYSADTLAHNNDIALKIRSKRGCAQPSRTIQTICL 300
Db 240 RSRLLNSTQCEMKFEVENLILHEDYSADTLAHNNDIALKIRSKRGCAQPSRTIQTICL 299
QY 301 PSWYNDPQ---FGTSCITGFGKNSDYLYPEQLKMTVVKLISHRECOQPHYGSVTT 357
Db 300 PSWYNDPNDPFPFSGCEITGFGKNSDYLYPEQLKMTVVKLISHRECOQPHYGSVTT 359
QY 358 KMLCAADPQKKTSCQDSSGGLVCSIQGHTLTGIVSWRGKALKDKPGVYTRVSHFLP 417
Db 360 KMLCAADPQWETBSCQDSSGGLVCSIQGHTLTGIVSWRGKALKDKPGVYTRVSHFLP 419
QY 418 WIRSHTEKENGAL 431
Db 420 WIRSHTEKENGAL 433

RESULT 3
UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
GN (U-plasminogen activator).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney; PubMed=6096832;
RX MEDLINE=85087954;
RT Negamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Negamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1db-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X01648; CA25806.1; -;
DR EMBL; X02724; CA26511.1; -;
DR PIR; A00932; UKPG.
DR HSSP; P00749; IKDU.
DR MEROPS; S01.231; -;
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR008293; Pept_S1A_uPA.
DR InterPro; IPR001254; Peptidase_S1.

```

```

DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PIRSF; PIRSF001144; Uro_kinase_act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle_1.
DR SMART; SM00130; KR_1.
DR SMART; SM00020; TRYSPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Plasminogen activator; Hydrolyase; Serine protease; Glycoprotein;
KW kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 KRINGLE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 190 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
FT ACT_SITE 241 241 Q -> H (IN REF. 1; CA25806).
FT CONFLICT 242 242 Q -> H (IN REF. 1; CA26511).
FT CONFLICT 288 288 A -> GS (IN REF. 1; CA25806).
SQ SEQUENCE 442 AA; 4916 MW; EE32FCEP50132EE CRC64;

Query Match 80.64; Score 1928.5; DB 1; Length 442;
Best Local Similarity 79.4%; Pred. No. 2,7e-145;
Matches 350; Conservative 33; Mismatches 47; Indels 11; Gaps 2;

```

QY 350 YVSEVTTKMLCAADPQWKTDSGCGSGGPLYCSLQGRMTITGVSWGRGALKDKPKGVY 409
 DB 361 YVSEVTTKMLCAADPQWKTDSGCGSGGPLYCSLQGRMTITGVSWGRGALKDKPKGVY 420
 QY 410 TRVSHFLPWIRSHTKENGGLA 430
 DB 421 TRVSRFLTWIHTHVGENGGLA 441

RESULT 4
 UROK_BOVIN STANDARD; PRT; 433 AA.
 ID UROK_BOVIN STANDARD; PRT; 433 AA.
 AC Q05389; Q28209;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 GN (U-plasminogen activator).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovine; Bos.
 NC NCB1_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aortic endothelium;
 RX MEDLINE=93216119; PubMed=8385052;
 RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
 RA Schleuning W.-D.;
 RT "Bovine urokinase-type plasminogen activator and its receptor:
 RT cloning and induction by retinoic acid.";
 RL Gene 125:177-183(1993).
 RN [2]

RP SEQUENCE OF 12-433 FROM N.A.
 RC TISSUE=Kidney;
 RA Ravn P., Berglund L., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen activators uPA
 RT and tPA.";
 RL Int. Dairy J. 5:605-617(1995).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- INDUCTION: By retinoic acid.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
 CC or send an email to license@ebi.ac.uk).

 CC EMBL: L03546; AAA51419.1; -
 DR EMBL: X85801; CAAS9796.1; -
 DR PIR: JN0560; JN0560.
 DR HSSP: P00749; ILMW.
 DR MEROPS: S01.231; -
 DR InterPro: IPR009003; Cys_ser_crypshn.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR008293; Pept_S1A_uPA.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00089; trypsin; 1.
 DR Pfam: PF00051; kringle; 1.
 DR PIRSF: PIRSF01144; Uro_k plasm act; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS00026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; signal; zymogen.
 FT SIGNAL 1
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
 FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 180 CONNECTING PEPTIDE.
 FT DOMAIN 181 433 SERINE PROTEASE.
 FT DISULFID 33 41 BY SIMILARITY.
 FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 211 227 BY SIMILARITY.
 FT DISULFID 219 290 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CONFLICT 189 189 A -> T (IN REF. 2).
 SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 76.6%; Score 1835; DB 1; Length 433;
 Best Local Similarity 75.1%; Pred. No. 6; ee-138;
 Matches 325; Conservative 46; Mismatches 60; Indels 2; Gaps 1;
 QY 1 MRALLRLILCVAVSPDSKSNELHGV--PSNCDCLNGGTCVSNKYSNTIMWCNCPKFG 58
 DB 1 MRVLLACLVALVSDSDSNEVHKESGSCGLNGKCVYKYSNIGRSCPKFG 60
 QY 59 GORCEIDSKTYEGNGHFFYRGASDTMGRPCLPMNSATVLOQTYAHRSALQLGLGK 118
 DB 61 GEHCEIDTSKTYQGNGHSYRGKANDLSGRPCLANDSPVLLKMYAHRSDAIQLGLGK 120
 QY 119 HNYCRNPNRRRWCTYQVGLKPLVQECVHDCADGKSSPPEBELFCGCGKTLRRFK 178
 DB 121 HNYCRNPNDRRWCTYQVGLKPLVQECVHDCADGKSSPPEBELFCGCGKTLRRFK 180
 QY 179 IIGGEFTTENOPWPAIYRRHGGSTVYCGSLISPCVVISATHCFTIDYPKKEDYIVY 238
 DB 181 IVGGQVTNANQWPAIYRRHGGSTVYCGSLISPCVVISATHCFTIDYPKKEDYIVY 240
 QY 239 IGRSRLNSNTGEMKEVENLILKDYSDTLAHNDIALLKRSKGRCAQPSRTIQT 298
 DB 241 LGSRSLNSDTRGEMQEVKLIHEDYSASLSLHNDIALLKRTSRGCAQPSRSIQT 300
 QY 299 CLPSTNDPQFGTSCETIGFGKENSVDYLYPELQKTVLKLISHRCQPHYIGSEVTTK 358
 DB 301 CLPPEHDAHRSRTCEITGFGKENSVDYLYPELQKTVLKLISHRCQPHYIGSEVTTK 360
 QY 359 MLCADPQWKTDSGCGSGGLVCSLQGRMTITGVSWGRGALKDKPKGVYTSVSHFLPW 418
 DB 361 MLCADPQWKTDSGCGSGGLVCSLQGRMTITGVSWGRGALKDKPKGVYTSVSHFLPW 420
 QY 419 IIRSHTKENGGLAL 431
 DB 421 INTHTRGEINLV 433

RESULT 5
 UROK_RAT

ID UROK_RAT STANDARD; PRT; 432 AA.
AC P29598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
GN (U-plasminogen activator).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer 344;
RX MEDLINE=92233409; PubMed=1568219;
RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
RA Kefford R.F.,
RT "Transcriptional and posttranscriptional activation of urokinase
RT plasminogen activator gene expression in metastatic tumor cells.";
RL Cancer Res. 52:2489-2496(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC Tissue=Kidney;
RA Rabbani S.A.;
RL Submitted (Apr-1992) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 156 in the low
CC molecular mass form to yield a short A1 chain (By similarity).
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63434; CAA45028.1; -;
DR EMBL; X65651; CAA46601.1; -;
DR PIR; S24604; S18932.
DR HSSP; P00749; 1KDU.
DR MEROPS; S01.231; -;
DR InterPro; IPR0009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR008293; Pept_S1A_uPA.
DR InterPro; IPR001254; peptidase_S1.
DR InterPro; IPR001314; peptidase_S1A.
DR Pfam; PF00051; Kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PIRSF; PIRSF001144; Uro_kinase; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN-
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF_1.
DR SMART; SM00130; KR_1.
DR SMART; SM00020; TRYD_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19
FT CHAIN 20 432
FT CHAIN 20 177
FT CHAIN 156 177
FT CHAIN 179 432
FT DOMAIN 27 63
FT DOMAIN 70 151
FT DOMAIN 152 178
FT DOMAIN 179 432
FT DISULFID 31 39
FT DISULFID 33 51
FT DISULFID 53 62
FT DISULFID 168 300
FT DISULFID 210 226
FT DISULFID 218 289
FT DISULFID 314 383
FT DISULFID 346 362
FT DISULFID 373 401
FT ACT_SITE 225 225
FT ACT_SITE 276 276
FT ACT_SITE 377 377
FT ACT_SITE 16 16
FT CONFLICT 24 24
FT CONFLICT 332 332
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;
Query Match 72.2%; Score 1728.5; DB 1; Length 432;
Best Local Similarity 71.3%; Pred. No. 1,8e-129;
Matches 308; Conservative 51; Mismatches 70; Indels 3; Gaps 3;

QY 1 MRALLRLLLCVLVVSDSKSNEH-QVPSNCDCLNGGCVSNKYSNIHMCCKPKFG 59
1 MRWLASLPLCAL-VANSEGSLEASDESNGCGVGVSVYKYSIRRCSPKPKFG 59
DB 1 MRWLASLPLCAL-VANSEGSLEASDESNGCGVGVSVYKYSIRRCSPKPKFG 59
QY 60 QHCEIDSKSTCYGNGHFYRGRKASTDTMGRPCIPMNSATVLOQTYHAHSDALQLGLGR 119
60 EHEEIDTSTKCYGNGQSYRGRKANTDKRPLCLANNSPAVLQOTYVAHRSDALSLGLGR 119
DB 60 EHEEIDTSTKCYGNGQSYRGRKANTDKRPLCLANNSPAVLQOTYVAHRSDALSLGLGR 119
QY 120 NTCRNDRNRPRPCVYQVGLKPLVORCMHDCADGKSPSPBELKFCGCGKTLRRFR 179
120 NTCRNDRNRPRPCVYQVGLKPLVORCMHDCADGKSPSPBELKFCGCGKTLRRFR 179
DB 120 NTCRNDRNRPRPCVYQVGLKPLVORCMHDCADGKSPSPBELKFCGCGKTLRRFR 179
QY 180 IGGEFTTIENQPFALYRRHGRGS-VTVYCGSLISPCWVISATHCFTIDYPKEDYIV 238
180 IGGEFTTIENQPFALYRRHGRGS-VTVYCGSLISPCWVISATHCFTIDYPKEDYIV 238
DB 180 VGEFETVENQPFALYRRHGRGS-VTVYCGSLISPCWVISATHCFTIDYPKEDYIV 238
QY 239 LGRSLNNTQGMKEVENLIIHKDYADTLAHNDIALLKIRSKRGCAOPSRITQTI 298
239 LGRSLNNTQGMKEVENLIIHKDYADTLAHNDIALLKIRSKRGCAOPSRITQTI 298
DB 239 LGRSLNNTQGMKEVENLIIHKDYADTLAHNDIALLKIRSKRGCAOPSRITQTI 298
QY 240 LQSKRNSYNPGEMKEVEQLIIHEDFSDETLAFHNDIALLKIRSTGCAOPSRITQTI 299
240 LQSKRNSYNPGEMKEVEQLIIHEDFSDETLAFHNDIALLKIRSTGCAOPSRITQTI 299
DB 240 LQSKRNSYNPGEMKEVEQLIIHEDFSDETLAFHNDIALLKIRSTGCAOPSRITQTI 299
QY 299 CLPSWYNDQFGFSCITGFGKENSFDVLYPBLQKMTVYKLSHRECOQPHYGSEVTK 358
299 CLPSWYNDQFGFSCITGFGKENSFDVLYPBLQKMTVYKLSHRECOQPHYGSEVTK 358
DB 299 CLPSWYNDQFGFSCITGFGKENSFDVLYPBLQKMTVYKLSHRECOQPHYGSEVTK 358
QY 300 CLPPRFGDAPFGSDCIITGFGSATDYFPYKOLKNSVVKIIISHECKOPHYGSEVTK 359
300 CLPPRFGDAPFGSDCIITGFGSATDYFPYKOLKNSVVKIIISHECKOPHYGSEVTK 359
DB 300 CLPPRFGDAPFGSDCIITGFGSATDYFPYKOLKNSVVKIIISHECKOPHYGSEVTK 359
QY 359 MLCADPQWKTDSCQDSGAPLVCSLQGRMTTLTGIVSMRGCALKDKPQVYTVSHFLPW 418
359 MLCADPQWKTDSCQDSGAPLVCSLQGRMTTLTGIVSMRGCALKDKPQVYTVSHFLPW 418
DB 359 MLCADPQWKTDSCQDSGAPLVCSLQGRMTTLTGIVSMRGCALKDKPQVYTVSHFLPW 418
QY 360 MLCADPEWKTDSCQDSGAPLVCSLQGRMTTLTGIVSMRGCALKDKPQVYTVSHFLPW 419
360 MLCADPEWKTDSCQDSGAPLVCSLQGRMTTLTGIVSMRGCALKDKPQVYTVSHFLPW 419
DB 360 MLCADPEWKTDSCQDSGAPLVCSLQGRMTTLTGIVSMRGCALKDKPQVYTVSHFLPW 419
QY 419 IHSHTKEENGLA 430
419 IHSHTKEENGLA 430
DB 419 IHSHTKEENGLA 430
420 IHSHTKEENGLA 431
420 IHSHTKEENGLA 431

RESULT 6
UROK_MOUSE STANDARD; PRT; 433 AA.
AC P06869;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

RX MEDLINE=90110185; PubMed=2295632;
 RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;
 RT "The chicken urokinase-type plasminogen activator gene.";
 RL J. Biol. Chem. 265:1339-1344(1990).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; J05187; AAA49131.1; -;
 DR EMBL; J05188; AAA49130.1; -;
 DR PIR; A35005; A35005.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.231; -;
 DR InterPro; IPR009003; Cys_ser_trypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR008293; Pept_S1A_uPA.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_STA.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYR_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Signal; Zymogen.
 FT SIGNAL 1 20
 FT CHAIN 21 434
 FT CHAIN 21 171
 FT CHAIN 173 434
 FT DOMAIN 36 72
 FT DOMAIN 79 158
 FT DOMAIN 159 172
 FT DOMAIN 173 434
 FT DISULFID 40 48
 FT DISULFID 42 60
 FT DISULFID 62 71
 FT DISULFID 162 296
 FT DISULFID 202 218
 FT DISULFID 210 285
 FT DISULFID 310 379
 FT DISULFID 342 358
 FT DISULFID 369 397
 FT ACT_SITE 217 217
 FT ACT_SITE 272 272
 FT ACT_SITE 373 373
 FT CARBOHYD 228 228
 SQ SEQUENCE 434 AA; 49400 MW; B0881048DD66A55 CQC64;

Best Local Similarity 46.0%; Pred. No. 1,5e-72;
 Matches 196; Conservative 66; Mismatches 134; Indels 30; Gaps 8;
 QY 10 LCVLVYS-DS-----KSGNELHQVPSNCDCLNGGCVANKKFSNTHMNCNCKPKKGGQH 61
 DB 11 LCTLVGLSDVYRYQYKLSHKRPOHRECCCLNGGICITIRFSPQIKRLCTCEGIGGLH 70
 QY 62 CETIDSKTCYEGNGHFYRGKASTDTMGRCPLPMSATVLO-QTYHAHRSDALQGLGKH 120
 DB 71 CEIDTNSICVSGNGEDYRGMAEDP-----GCLYMDHPSVIRMGDYHADLKNALQGLGKH 126
 QY 121 YCNPNDRRRPPCYVYGLKPLVQECMADCDGKSPSPPELKKQCCQKTLRPRKTI 180
 DB 127 YCNPNDRRRPPCYVYGLKPLVQECMADCDGKSPSPPELKKQCCQKTLRPRKTI 174
 QY 181 GGFETTENQPMFAIYRRRGGSVTVYVCGSGLISPCWVISATHCFID----YPKEDYI 236
 DB 175 GGSQAEVETQPMJAGIFQNTM-GTDQFLCGSGLIDPCWVLTAAHCFNPKPKQPNKSVYK 233
 QY 237 VYLGRSLNSNTQGEKMEFEVENLILKDYASADTLAHNNDIALIKIRSKGRCAQPSRTIQ 296
 DB 234 VFLGKSLTNDDEHQFVWDEILSHDPFDHTGNDNDIALIRIRASGCAVESNYVR 293
 QY 297 TILCPMTNDPQGTSCETIFGKENSIDYLYPEQLKMTYVYKLSHRECCQPHYGSSEVT 356
 DB 294 TVCLPCKNLMLYDNTWCEIINGKONSIDYLYAQRILMSATVNLISODCKNKYDSTRTV 353
 QY 357 TKMLCAADPQWKTDSCGSGGPLYCSLOGRMILTGVISVGRGALCKDPGYTVRVSHFL 416
 DB 354 DNWVCADPLMETDACKGDSGGPVCENHGRMTLYGLVSGDGCANKNKGVITRYVRYL 413
 QY 417 PWIRSH 422
 DB 414 NWIDSN 419
 RESULT 8
 ID URT2 DESRO STANDARD; PRT; 477 AA.
 RT URT2 DESRO
 AC P15638;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
 DE alpha-2) (BAT-PA) (T-plasminogen activator).
 OS Desmodus rotundus (Vampire bat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 CC Desmodontinae; Desmodus.
 OX NCBI_Taxid=9430;
 RN 1
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN 2
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Salivary gland;
 RX MEDLINE=90036867; PubMed=2509450;
 RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
 RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
 RT "Isolation, characterization, and cDNA cloning of a vampire bat
 RT salivary plasminogen activator.";
 RL J. Biol. Chem. 264:117947-117952(1989).
 RN 3
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,

Query Match

42.1%; Score 1009; DB 1; Length 434;

RA Donner P.;
 RT "plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -1- FUNCTION: Probably essential to support the feeding habits of this exclusively haematophagous animal. Probable potent thrombolytic agent.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
 CC -1- ENZYME REGULATION: Activity toward plasminogen is stimulated in the presence of fibrin I.
 CC -1- SUBUNIT: Monomer.
 CC -1- DOMAIN: The fibronectin type-I domain mediates binding to fibrin, and the kringle domain apparently mediates fibrin-induced stimulation of activity.
 CC -1- SIMILARITY: Belongs to peptide family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL: M63988; AAA31593.1; -;
 DR EMBL: J05082; AAA31596.1; -;
 DR PIR: A34369; A34369.
 DR PIR: J05098; J05098.
 DR HSSP: P98119; IAS1.
 DR MEROPS: S01.232; -;
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR006209; EGF_Like.
 DR InterPro: IPR006083; Fibronctn.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR006001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00039; fn1; 1.
 DR Pfam: PF00051; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF_1.
 DR SMART: SM00130; FN1; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS50026; EGF_3; 1.
 DR PROSITE: PS01253; FIBRONECTIN_1; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS50070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Plasminogen activation: Hydrolyase; Serine protease; Glycoprotein; KM
 Kringle; EGF-like domain; Signal; Multigene family.
 FT STGNL 1 36
 FT CHAIN 37 477
 FT DOMAIN 40 82
 FT DOMAIN 83 121
 FT DOMAIN 128 209
 FT DOMAIN 225 477
 FT ACT_SITE 225 272
 FT ACT_SITE 321 321
 FT ACT_SITE 428 428
 FT DISULFID 42 72
 BY SIMILARITY.

FT DISULFID 70 79 BY SIMILARITY.
 FT DISULFID 87 98 BY SIMILARITY.
 FT DISULFID 92 109 BY SIMILARITY.
 FT DISULFID 111 120 BY SIMILARITY.
 FT DISULFID 128 209 BY SIMILARITY.
 FT DISULFID 149 191 BY SIMILARITY.
 FT DISULFID 180 204 BY SIMILARITY.
 FT DISULFID 214 345 BY SIMILARITY.
 FT DISULFID 257 273 BY SIMILARITY.
 FT DISULFID 265 334 BY SIMILARITY.
 FT DISULFID 359 434 BY SIMILARITY.
 FT DISULFID 391 407 BY SIMILARITY.
 FT DISULFID 424 452 BY SIMILARITY.
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 403 403 Y -> K (IN REF. 2).
 FT CONFLICT 417 417 Y -> H (IN REF. 2).
 FT CONFLICT 435 435 M -> R (IN REF. 2).
 FT SEQUENCE 477 AA; 53719 MW; 17486555085077C CRC64;
 SQ
 Query Match 36.3%; Score 868.5; DB 1; Length 477;
 Best Local Similarity 42.7%; Pred. No. 2,2e-61;
 Matches 179; Conservative 61; Mismatches 150; Indels 29; Gaps 9;
 QY 19 KGSNELHQP-----SNCDLNGGTGVSNKXPSNIHMCNPKFGGCHCEIDSKTYEGN 74
 DB RGLAQCHTVVKSCELRGNGGTQWASFSDF-VCQCPKGYTGKCEVDTHATCYKQ 132
 QY 75 GHFYRGASTDMGRCLPNSATVLAQYTHARSDALGLKKNYCRPNRRRWCV 134
 DB 133 GVTYRGWSTSESAGACINNSMLTRRYNGRSDITGLGNHNCRPNDDNSKWCY 192
 QY 135 VQVGLKPLVQECVHDCADGKKSPPEELKPGCG-QKTLRPFKLIGSEFTIENQWF 193
 DB 193 VTRASKFILEFCVPCPS-----KATGLARKYEPQLHSGGLFTDITSPWQ 240
 QY 194 AAIYRRHRRGS-VTYVCGGSLSPCWISATHCFID-YPKEDYIYLGSRINSNTQGE 251
 DB 241 AALFAQNRSSGRFLCGGILISSCWVLAHCFQERYPPQHLRV-IGRTYRVKPKER 299
 QY 252 MKREVENLIHKYASDTLAHNDIALKTRSEGRCAQPSRTIQTCLEPSMNDPPQGT 311
 DB 300 QTFEVEKCIYHEBDDT--YNNDIALLQKSSPCQAOBDSVRALCLPEANLQLPDWT 357
 QY 312 SCEITRGKENSIDYLYPEQLKMTVVKLSHRCCQPHYVYGEVTTKMLCAADPQMT-- 369
 DB 358 ECELSTGYGKHKSSPYSBQLKGVRLYPSRCTSKFLPKVTYNNMLCAQDTRSGEIT 417
 QY 370 ----DSQCGDGGPLVCSLOGRMTLTGIYSWGRGALKDKPGVYTRVSHFLPMIRSHTK 424
 DB 418 PNVDACQDGGSGPLVCMNDNHTLTGIIISWVGCGEKOIPGYVTVYTLGMIIDMR 476
 RESULT 9
 ID TPA HUMAN STANDARD; PRT; 562 AA.
 AC P00750; O15103;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
 DE (t-PA) (t-plasminogen activator) (Alteplase) (Retepase).
 GN PLAT.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Melanoma;
 RX MEDLINE=83115262; PubMed=6337343;
 RA Pennica D., Holmes W.E., Kohn W.J., Harkins R.N., Vohar G.A.,
 Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,

RA Goeddel D.V., Collen D.;
 RT "Cloning and expression of human tissue-type plasminogen activator
 RT cDNA in E. coli.";
 RT Nature 301:214-221(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal lung;
 RX MEDLINE=88262579; PubMed=313640;
 RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
 RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
 RT from human fetal lung cells.";
 RL Nucleic Acids Res. 16:5695-5695(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88054470; PubMed=2824147;
 RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Wackins P., Galli J.,
 RA Heising N.;
 RT "Expression of human uterine tissue-type plasminogen activator in
 RT mouse cells using BPV vectors.";
 RL DNA 6:461-472(1987).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86196143; PubMed=3009482;
 RA Frieze Degen S.J., Rajput B., Reich E.;
 RT "The human tissue plasminogen activator gene.";
 RL J. Biol. Chem. 261:6972-6985(1986).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84298137; PubMed=6089198;
 RA NY T., Elgh F., Lund B.;
 RT "The structure of the human tissue-type plasminogen activator gene:
 RT correlation of intron and exon structures to functional and
 RT structural domains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86284200; PubMed=3090401;
 RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
 RA Opdenacker G., Volckaert G., Kombeuts W., Billiau A., Somer P.;
 RT "Cloning of cDNA coding for human tissue-type plasminogen activator
 RT and its expression in Escherichia coli.";
 RL Mol. Biol. Med. 3:279-292(1986).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOPORM SHORT).
 RC TISSUE=Umbilical vein;
 RX MEDLINE=90192129; PubMed=2107528;
 RA Siebert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells.";
 RL Nucleic Acids Res. 18:1086-1086(1990).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strussberg R.L., Feingold R.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.T., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
 RA Stalestein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rabe S.S., Lonnellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley R.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smalios D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maitra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [9]
 RP SEQUENCE OF 212-361 FROM N.A.
 RX MEDLINE=83169656; PubMed=6572897;
 RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
 RA Josephson S.;
 RT "Isolation of cDNA sequences coding for a part of human tissue
 RT plasminogen activator.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
 RN [10]
 RP SEQUENCE OF 1-36 FROM N.A.
 RX MEDLINE=85289338; PubMed=3161893;
 RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
 RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region.";
 RL J. Biol. Chem. 260:11223-11230(1985).
 RN [11]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1368681;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
 RL Agric. Biol. Chem. 55:1225-1232(1991).
 RN [12]
 RP SEQUENCE OF 36-562.
 RC TISSUE=Melanoma;
 RX MEDLINE=85000468; PubMed=6433976;
 RA Pohl G., Kaelin-Lang A., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences.";
 RL Biochemistry 23:3701-3707(1984).
 RN [13]
 RP SEQUENCE OF 33-52 AND 311-330.
 RC TISSUE=Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator.";
 RL Eur. J. Biochem. 132:681-686(1983).
 RN [14]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells.";
 RL Eur. J. Biochem. 186:273-286(1989).
 RN [15]
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 RX MEDLINE=9159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain.";
 RL Biochemistry 30:2311-2314(1991).
 RN [16]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=91244765; PubMed=1645336;
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaaskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in Escherichia coli.";
 RL J. Biol. Chem. 266:10070-10072(1991).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Landa D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RA Bode W.;
 RT "The 2.3 Å crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator.";
 RL J. Mol. Biol. 258:117-135(1996).
 RN [18]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;


```

DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS0186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1
FT CHAIN 37
FT DOMAIN 37 431
FT DOMAIN 82 163
FT ACT_SITE 179 431
FT ACT_SITE 275 226
FT ACT_SITE 382 275
FT DISULFID 41 52
FT DISULFID 46 63
FT DISULFID 65 74
FT DISULFID 82 163
FT DISULFID 103 145
FT DISULFID 134 158
FT DISULFID 168 299
FT DISULFID 211 227
FT DISULFID 219 288
FT DISULFID 313 388
FT DISULFID 345 361
FT DISULFID 378 406
FT CARBOHYD 139 139
FT CARBOHYD 352 352
SQ SEQUENCE 431 AA; 48221 MW; 699BSB675B162CBF CMC64;

Query March 36.2%; Score 867.5; DB 1; Length 431;
Best Local Similarity 41.6%; Pred. No. 2.4e-61;
Matches 185; Conservative 61; Mismatches 152; Indels 47; Gaps 10;

4 LLARLLLCVLVVSQK-----GSNELHVPSCDCLNGTGVSKKYSNI 48
9 LLCVLLLCGAVFSLPQRTYRQLARGSAVGCSELR-----CFNGGTCWQASPSDF 61
49 HWCNCPKPKGGQCEIDKSKTCYEGNGHFYRGKASTDTMRPCLPMNSATVLAQTYAHR 108
62 -VCCCPKGYTGKQCEVDTHATCYKQGVYRGWTSTBSGAQCINMNSNLLTRITNGRR 120
109 SDALOLGIGKKNYGRNPNRRRPPCYQVGLKPLVQECMHDCADGKKPSPPEELKQC 168
121 SDATLLGNNHYCRNPNNSKPPCYVAKSKFTLEFCSPVVC-----KATC 168
169 G-QKTLRPFKILIGEFTTLENQWPAFYRHRGS-VTVVCGSLISPCWVISATHCF 226
169 GLRKYRKPQLHSTGSLFTDISHWQAAIFQKNRSGSERLCCGILLISCVUTAHNCF 228
227 ID-YPKEDYIVYGRSLNSNTGEMKEFEVENILHKDVSADTLAHDNIALKIRSK 285
229 QERYVPQHLRVV-LGRYRVKPGKEQTFEVEKCIHHEFDPT--VNNIDALLQLKSGS 285
286 GRCAQPSRTIQTICLPQMYNDPQGTSCETLPGKKNSTDYLYEOLKMTVVVLLISREC 345
286 PQCAQESDSVAATCLPEANILQLPWTECELSGYKHKSSPFYSBQLKEGIVRLYPSRC 345
346 QQPHYVSEVTTKMLCADPQMT-----DSCGDSGGPLVCSLQGRMTLTGIVWGRG 399
346 TSKRLFKVTYVNNMLCGADTRSGEIIYPNVHACGCGDSGGLVVCNNDNMHTLLGIISVG 405
400 CALKQKFGVYTRVSHFLPWIRSHTK 424
406 CGEKDIPGVYTKVNYLGIWIRDNRR 430

```

```

TPA_RAT
ID TPA_RAT STANDARD; PRT; 559 AA.
AC P19637;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (cPA)
DE (c-PA) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9170114; PubMed=3148445;
RA My T., Leonardson G., Heuvel A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator."
RL DNA 7:671-677(1988).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90130448; PubMed=2105315;
RA Feng P., Ohlsson M., Ny T.;
RT "The structure of the TATA-less rat tissue-type plasminogen activator
RT gene. Species-specific sequence divergences in the promoter predict
RT differences in regulation of gene expression."
RL J. Biol. Chem. 265:2022-2027(1990).
CC -1- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
CC controlling plasmin-mediated proteolysis, it plays an important
CC role in tissue remodeling and degradation, in cell migration and
CC many other physiological events.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
CC bond.
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: Binds to the kringle structure of the fibrin A
CC chain. Binding to fibrin enhances its catalytic activity.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 2 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M23697; AAA41812.1; -
DR EMBL; M31197; AAA42261.1; -
DR EMBL; M31185; AAA42261.1; JOINED.
DR EMBL; M31186; AAA42261.1; JOINED.
DR EMBL; M31187; AAA42261.1; JOINED.
DR EMBL; M31188; AAA42261.1; JOINED.
DR EMBL; M31189; AAA42261.1; JOINED.
DR EMBL; M31190; AAA42261.1; JOINED.
DR EMBL; M31191; AAA42261.1; JOINED.
DR EMBL; M31192; AAA42261.1; JOINED.
DR EMBL; M31193; AAA42261.1; JOINED.
DR EMBL; M31194; AAA42261.1; JOINED.
DR EMBL; M31195; AAA42261.1; JOINED.
DR EMBL; M31196; AAA42261.1; JOINED.
DR EMBL; A19618; CA01482.1; -
DR PIR; A35029; A35029.
DR HSSP; P00750; IRTF.

```

DR	MEROPE:	:501..232;	-	Cys Ser trypsin.
DR	InterPro:	IPR009003;		Egf_like.
DR	InterPro:	IPR006209;		Egf_like.
DR	InterPro:	IPR000083;		Fibrinctn.
DR	InterPro:	IPR006210;		IgGF.
DR	InterPro:	IPR000001;		Kring1e.
DR	InterPro:	IPR001254;		Peptidase S1.
DR	InterPro:	IPR001314;		Peptidase_S1A.
DR	Pfam:	PF00008	:	Egf; 1.
DR	Pfam:	PF00039	:	fml; 1.
DR	Pfam:	PF00051	:	kringle; 2.
DR	Pfam:	PF00089	:	trypsin; 1.
DR	PRINTS:	PR00722;		CHYMOTRYPSIN.
DR	PRINTS:	PR00018;		KRINGLE.
DR	ProDom:	PD000395;		kringle; 2.
DR	SMART:	SMO0181;		Egf; 1.
DR	SMART:	SMO0058;		fml; 1.
DR	SMART:	SMO0130;		Kr; 2.
DR	SMART:	SMO0020;		TRYP_Spc; 1.
DR	PROSITE:	PS00022;		Egf_1; 1.
DR	PROSITE:	PS01186;		Egf_2; 1.
DR	PROSITE:	PS50026;		Egf_3; 1.
DR	PROSITE:	PS01253;		FIBRONECTIN_1; 1.
DR	PROSITE:	PS00021;		KRINGLE_1; 2.
DR	PROSITE:	PS50070;		KRINGLE_2; 2.
DR	PROSITE:	PS50240;		TRYPSIN_DOM_1.
DR	PROSITE:	PS00134;		TRYPSIN_His; 1.
DR	PROSITE:	PS00135;		TRYPSIN_SER; 1.
KM	Plasma:	kinase activation:		Hydrolase; Serine protease; Glycoprotein;
FT	SIGNAL	1	17	PROBABE.
FT	PROPEP	18	29	
FT	CHAIN	30	559	TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT	CHAIN	30	308	TISSUE-TYPE PLASMINOGEN ACTIVATOR A CHAIN.
FT	CHAIN	309	559	TISSUE-TYPE PLASMINOGEN ACTIVATOR B CHAIN.
FT	DOMAIN	36	78	FIBRONECTIN TYPE-1.
FT	DOMAIN	79	117	Egf-like.
FT	DOMAIN	124	205	KRINGLE 1.
FT	DOMAIN	213	294	KRINGLE 2.
FT	DOMAIN	309	559	SERINE PROTEASE.
FT	ACT_SITE	355	355	CHARGE RELAY SYSTEM.
FT	ACT_SITE	404	404	CHARGE RELAY SYSTEM.
FT	ACT_SITE	510	510	CHARGE RELAY SYSTEM.
FT	DISELID	38	68	BY SIMILARITY.
FT	DISELID	66	75	BY SIMILARITY.
FT	DISELID	83	94	BY SIMILARITY.
FT	DISELID	88	105	BY SIMILARITY.
FT	DISELID	107	116	BY SIMILARITY.
FT	DISELID	124	205	BY SIMILARITY.
FT	DISELID	145	187	BY SIMILARITY.
FT	DISELID	176	200	BY SIMILARITY.
FT	DISELID	213	294	BY SIMILARITY.
FT	DISELID	234	276	BY SIMILARITY.
FT	DISELID	265	289	BY SIMILARITY.
FT	DISELID	297	428	INTERCHAIN (by similarity).
FT	DISELID	340	356	BY SIMILARITY.
FT	DISELID	348	417	BY SIMILARITY.
FT	DISELID	442	516	BY SIMILARITY.
FT	DISELID	474	490	BY SIMILARITY.
FT	DISELID	506	534	BY SIMILARITY.
FT	CARBOHYD	149	149	N-LINKED (GLCNAC...)
FT	CARBOHYD	481	481	N-LINKED (GLCNAC...)
FT	CONFLICT	380	380	E -> K (IN REF. 1).
QO	SEQUENCE	559 AA;	62903 MM;	7DBD809C1DC921 CRC64;

Db	69	NSGLAQGSVPRVSCSBRFCNFGCCQALVPSDF--VCQCPDGFVGRCDIDTRATCFEG	127
Qy	74	NGHFPRGASIDITMRPCLPNNASVAVLQOTYAHNSDALQLGIGKNNYCRNPNNRRPWC	133
Db	128	QGITTRGWTSAENAGCAECINNNSSALSKQPYARRPNAIKLGIGNNYCRNPDRVPCWC	167
Qy	134	VYQVGLKFLVQECM-----VHDCADQK-----KPSPP-----	161
Db	188	YVFAKGIKTTTEFCSTIPACPKGPTEDCYGKVTVYRGTHSFTTSSKASCLFNNSMILGKTY	247
Qy	162	-----BELKFO-----CGQKTLR--PRF	177
Db	248	TAMRANSQALGIGRNNYCRNPDDGAKPMCHVKDKRLTWETCYDMSPCSTCGLRQYKQPGF	307
Qy	178	KIIGGEFTTINQPPFAIY--RRHGGSVTVYCGSGLSPCVAVISATHCFIDYPKKEDI	226
Db	308	RIKGGLFTDITSHPMQAAIFVNNKRSRPGRRFLCGVLISCVLSSAHCFFVEFRPHHLK	367
Qy	237	VYLGSRRLNSTQGMKFEVENTLIHKQYSAPTLHAHNDIALLKTRSKGRCQAPRTIQ	296
Db	368	VVLGRTYVNVPEEESQTEIEIKYIVHKEFDDT--YNDIDIALQLRSDSQCAQESSVVG	425
Qy	297	TICLPSMTNDPOF-----STSCETTFGKENSVDYLYPEQLKMTVYGLISHRECQPHYYG	352
Db	426	TACLP-----DDVQVLPDMWTBECLESTGYGHEASPPFSDBLKEAHRLVYSSRSTQHLFN	481
Qy	353	SEVTTKMLCAADP-----QWKTDSCQGSBGFLVCSLOGRMVTLTGIVSGRCALXDKPG	407
Db	482	KTITSNMLCAGDTRTGQNDVHDAQGSBGFLVCMIDKRMVTLGIISWGLCGCGQDVPG	541
Qy	408	VYTRVSHFLPWTRSHKE	425
Db	542	YTKVTNYLNMVQDNMKQ	559

Query Match	Best Local Similarity	Score 841.5; DB 1; Length 477;
Matches 176; Conservative 57; Mismatches 151; Indels 37; Gaps 10.		
19 KGSNELHOVPSN-CD---CLNGGTCVSNKYFSPNIIHWCNCPKPKFGQGHCEIDSKTCYGN 74	92 109	BY SIMILARITY.
74 RGQARCHTVPVNSCSBPCFNGGTQWAVYFSDF-VCCCPAGYTGRCRCVDTRATCYEQ 132	111 120	BY SIMILARITY.
75 GHYYRKGASIDTDMGRCPCLPNSATVLCQTYHHRSDALQLGKXNYCENPNRRRPMCY 134	128 209	BY SIMILARITY.
133 GVYYRGRTWSTAESRVECCINWSSLTFRITYNGMPDAPFLGIGNINYCNPNRGAPKPMCY 192	149 191	BY SIMILARITY.
135 VVGVLKPLVQECVHDCADGKKSPPEELKQCQ-QKTLRPFKIIGSEFTTINQPMF 193	180 204	BY SIMILARITY.
	214 345	
	257 273	
	265 334	
	359 434	
	391 407	
	424 452	
	153 153	
	398 398	
	214 215	
	223 224	
	226 227	
	230 231	
	234 236	
	238 239	
	240 245	
	254 263	
	264 265	
	266 269	
	271 273	
	280 282	
	284 287	
	297 298	
	300 309	
	311 312	
	315 317	
	319 320	
	323 328	
	338 338	
	339 340	
	341 341	
	345 345	
	349 350	
	355 356	
	358 363	
	366 366	
	374 374	
	379 385	
	388 390	
	393 398	
	403 404	
	405 409	
	425 426	
	428 429	
	431 436	
	437 438	
	439 448	
	455 456	
	459 463	
	464 467	
	468 474	
	53616 MW; AA06FD1739C10E5E CRC64;	


```

Db 193 VIRAGFTSSCSVPVCS-----KATGLRKKEPOLHSTGGLFTDITSHPMQ 240
QY 194 AAIYRRHRRGGS-VYVCGGSLISPCWVISAHPICFDYPKEDYI-----VYLGSRNSN 247
Db 241 AAIIPAQRBRSSGGERFLCGILLISGCVLTAAHC-----GESYLPDLKVLGRTYVXP 295
QY 248 TQGBMKFEVENLIHKDYSADTLAHNDIALKIRSRKGCAPSRITQITCLPSWYNDP 307
Db 296 GSEBQTFKVKYIVHKEFDDT--YNNDIALLQKSDSPQCAQSDSVAICLPEANLQL 353
QY 308 QFETSCITTFGKENSNDIYLPBOLKMTYVKLISHRECCQPHYGVSEVTKMLCAADPQW 367
Db 354 PDWTECLLSYGRKSSSPFSEQLKKGHVLPSSRCAPKFLPKYVNNMCAQDTRS 413
QY 368 KT-----DSCQGDSPGGLVCSLQGRMTLGIYWSGRCALCKPKGYVTRVSHPLPIRS 421
Db 414 GEIYPNVHDCQGDSSGGLVCMNDNMTLGIISMGVCGEKDVPYVTVNTLGMIRD 473
QY 422 H 422
Db 474 N 474

RESULT 13
TPA_MOUSE STANDARD; PRT; 559 AA.
AC P1214; 091VP2;
DT 01-JUL-1989 (Rel. 11, Created)
DT 10-OCT-2003 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tpa)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8087303; PubMed=2826484;
RA Rickles R.J., Barrow A.L., Strickland S.;
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT activator mRNA and its expression during F9 teratocarcinoma cell
RT differentiation."
RL J. Biol. Chem. 263:1563-1569(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Datchenko L., Matulis K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Raha S.S., Loughell N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R.M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
CC controlling plasmin-mediated proteolysis, it plays an important
CC role in tissue remodeling and degradation, in cell migration and

```

```

CC many other physiopathological events.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
CC bond.
CC -!- SUBCELLULAR LOCATION: Secreted, extracellular.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-108 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: Binds to the kringle structure of the fibrin A
CC chain. Binding to fibrin enhances its catalytic activity.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03520; AAA40470.1; -.
DR EMBL; BC011256; AAH11256.1; -.
DR PIR; A29941; A29941.
DR HSP; P00750; 1ASH.
DR MEROPS; S01.232; -.
DR K0D; MGI:97610; Plat.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006083; Fibronctn.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1_1.
DR Pfam; PF00051; kringle_2.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYOTRYPsin.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle_2.
DR SMART; SM00181; EGF_1.
DR SMART; SM00058; FN1_1.
DR SMART; SM00130; KR_2.
DR SMART; SM00020; TRYP_SPC_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 359
FT ACT_SITE 355 355
FT ACT_SITE 404 404
FT ACT_SITE 404 404

```

```
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 G -> A (IN REF. 1).
FT CONFLICT 260 260 P -> A (IN REF. 1).
FT CONFLICT 325 325
SQ SEQUENCE 559 AA; 63122 MM; 8CCE2BD94514D9 CRC64;

Query Match 34.6%; Score 828.5; DB 1; Length 559;
Best Local Similarity 35.5%; Pred. No. 3.9e-58;
Matches 177; Conservative 66; Mismatches 158; Indels 97; Gaps 10;

18 SKGSNELHOVP-----SNCDINGSTCVSNKYFSNIHWCNPKKFGGCHIDSKTCYEG 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 NSGLVQCHSVPRSCSEPCRCNGGTCQALYFSDP--VCQCPDGHVGRKCDIDTRATCFEE 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
74 NGHYRGAASLDITMGRPCLPMNSATVLOQTHARSDLOGLGKHYCRPNDRRRPWC 133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
128 QGITRRGWSYAESGAEICINNSVLSLKPYNARPNAIKLGANHYCRPNDRDLKPMC 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
134 YVVGGLKPLVOECWHDCAQDKKPS----- 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
188 YVFRAGKTTTEFCSTPACPKKSEBCYVGKGYRGTGHSLTTSQASCLPMNSIVLMGKSY 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
159 -----SPPEELK-----FOCG-QKTLRPRF 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
248 TAWRTNSQALGLGRHNYCRNPDGARPMCHWKDKRLTWEXCDMSPCSTCGLRQYKRPQF 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
178 KIIGEFETINQPFPAIY--RRHGGSVTVVCGSLISPCMVISATCFIDYRKEDI 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
308 RIKGGLYDITSHPWQAPLIFYNNKRSPEGRFLCGSVLISSCVLSAHCFLERRPPNHLK 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
237 VYLGSRINSTQGEKKEVENLILHKDYADTLAHNDIALILKRSKRGCAOPSRITQ 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
368 VVLGRTYAVPGESEOTFEIKYIHEHFDLT--YDNDIALLOLRSGSKCAQSSSVG 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
297 TICLPSTMYNDPQF-----GTSCETITFGKENSTDYLYPPQLKMTYVKLISHRECOQPHYG 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
426 TACLP-----DPMQLPDWTECELSGCGKHEASPPFSRLKKAHYRLVPSRSCTSOHIFN 481
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
353 SEVTTKMLCAADP-----QMKTDSCQSGSGGLVOSLQGRMTLGIIVSMGRCALKDXPG 407
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
482 KTVINNMICAGDTRSGGNQDLHDAQQSGGGLVLCINKOMTLITGISNGLCGGKQADYVG 541
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
408 VYTRVSHFLPWIRSHSTKE 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
542 VYTKVTNYLDWTHDMKQ 559
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
TPA_BOVIN STANDARD; PRT; 566 AA.
ID_TPA_BOVIN
AC 028198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (TPA)
```

```
DE (t-PA) (t-Plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteleostomi; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA.";
RL Int. Dairy J. 5:605-617(1995).
CC - FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
controlling plasmin-mediated proteolysis, it plays an important
role in tissue remodeling and degradation, in cell migration and
many other physiological events.
CC - CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC - SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
bond.
CC - SUBCELLULAR LOCATION: Secreted, extracellular.
CC - PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-314 CATALYZED BY PLASMIN, TISSUE KALIKREIN OR FACTOR XA.
CC - MISCELLANEOUS: Binds to the kringle structure of the fibrin A
chain. Binding to fibrin enhances its catalytic activity.
CC - SIMILARITY: Belongs to peptidase family S1.
CC - SIMILARITY: Contains 1 EGF-like domain.
CC - SIMILARITY: Contains 1 fibronectin type I domain.
CC - SIMILARITY: Contains 2 kringle domains.
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X85800; CA59795.1; -.
DR HSSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006083; FibFnctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF. 1.
DR Pfam; PF00039; fn1. 1.
DR Pfam; PF00051; kringle. 2.
DR Pfam; PF00089; trypsin. 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle. 2.
DR SMART; SMO0181; EGF. 1.
DR SMART; SMO0058; FN1. 1.
DR SMART; SMO0130; KR. 2.
DR SMART; SMO0020; TRYD_SPC. 1.
DR PROSITE; PS00022; EGF_1. 1.
DR PROSITE; PS01186; EGF_2. 1.
DR PROSITE; PS50026; EGF_3. 1.
DR PROSITE; PS01253; FIBRONECTIN_1. 1.
DR PROSITE; PS00021; KRINGLE_1. 1.
DR PROSITE; PS50070; KRINGLE_2. 2.
DR PROSITE; PS50240; TRYPSIN_DOM. 1.
DR PROSITE; PS00134; TRYPSIN_HIS. 1.
DR PROSITE; PS00135; TRYPSIN_SER. 1.
DE Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
```

Query Match	34.3%	Score 821	DB 1	Length 566
Best Local Similarity	36.4%	Pred. No. 1.5e-57		
Matches 180	Conservative 62	Mismatches 160	Indels 92	Gaps 11
OY	20	GSNELHQP-SNCD---CLNGGTVCVSNKYFNSNIHMCNCPKFGGQHCIEIDKSTCYEGNG	75	
Db	75	GRAQCHSPVAKSCSEPCWPCFNGTCTCALYSNP-VCQCEEGMGKICEIDATRTCYKDG	133	
OY	76	HFYRGKASTDTGRDCLPWN SATVLQOTYHAHRSDALQGLKKNYCNPNDRRRRCVY	135	
Db	134	VAYRTGWSIAEGAGACAMNNSGLAMKPYSGRRPNAIRGLGNHYCRAPDQSKWCVY	193	
OY	136	QVGLKPLVQECVHNDCA---DGK-----	155	
Db	194	FKAGKYISFCSTPAACAKVAEBDDCYTGNGLAYRTRSHTSGASCLEPWN SVFLTSKIY	253	
OY	156	---KSSP-----PEELKTRQ-----CG-QKTLRPF	177	
Db	254	TAWKSNAPRALGLKNNHCNPDGDAQPMCHVWKDRQLTWEYCDVPQCVTGGRLQYRPPY	313	
OY	178	KIIGSEFTTIENQPMFAIY-RRHRGGSVTVYVGGSLISPCWVSIATHCIFDYPKKEDYI	236	
Db	314	RIKGGELFADTISHPMQALFYVKNRRSGRGRFLCGGLISSCWVLAHCFQERYPPHHLK	373	
OY	237	VYLGSRLSNSTOGEMKEVENLILHKDYSADTLAHNDIALIKIRSKGRCAQPSRTIQ	296	
Db	374	VFLGRTYRLVPEEBOGTPEVEKYIIHKPEDDDT--YNDNIALHLHLSDELTCARBSASVR	431	
OY	297	TICLPSMTNDPFGISCEITGFGENKSTIYLYPEOLAKTVVYKLISIREQOQPIYYSFT	356	
Db	432	TICLPDASIQLPDMTECEELSGYKHESSSPFSEBRLEKHAHVRLYPSSRCTSQHLFRVTY	491	
OY	357	TKMLCAADPQW-----KTDSGCGSGGGLVCSLQGRMTLTGVISKRGALCDKRGVYT	410	
Db	492	NNMLCAGDTRSGGDHTNLHDACQAGSGGGLVCKKDNHMLTVGLIISNGLGCGRKDVPGVTT	551	

```

OY      411 RVSHFLPWIRSHTK 424
       :::: |||:::
Db      552 KVTNYLDWIRDNTR 565

RESULT 15
ID      URTG DESRO      STANDARD;      PRT;      394 AA.
AC      P49150;
DT      01-FEB-1996 (Rel. 33, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA
DE      gamma).
OS      Desmodus rotundus (Vampire bat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC      Desmodontinae; Desmodus.
CX      NCBI_TaxId=9430;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Salivary gland;
RA      MEDLINE=92039036; PubMed=1937019;
RA      Kretzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA      Alagon A., Donner P., Schleuning W.D.;
RT      "The plasminogen activator family from the salivary gland of the
RT      vampire bat Desmodus rotundus: cloning and expression.";
RL      Gene 105:229-237(1991).
RN      [2]
RP      CHARACTERIZATION.
RA      MEDLINE=93393059; PubMed=1309059;
RA      Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA      Kretzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
RA      Donner P.;
RT      "Plasminogen activators from the saliva of Desmodus rotundus (common
RT      vampire bat): unique fibrin specificity.";
RL      Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC      -1- FUNCTION: Probably essential to support the feeding habits of this
CC      exclusively haematophagous animal. Probable potent thrombolytic
CC      agent.
CC      -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC      plasminogen to form plasmin.
CC      -1- SUBUNIT: Monomer.
CC      -1- SIMILARITY: Belongs to peptidase family S1.
CC      -1- SIMILARITY: Contains 1 kringle domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics, and the EMBL outstation at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M63990; AAA31595.1; -.
DR      PIR; J50600; J50600.
DR      HSSP; P98119; 1A51.
DR      MEROPS; S01.239; -.
DR      InterPro; IPR009003; Cys_Ser_trypsin.
DR      InterPro; IPR000001; Kringle.
DR      InterPro; IPR001254; Peptidase_S1.
DR      InterPro; IPR001314; Peptidase_S1A.
DR      Pfam; PF00051; kringle_1.
DR      Pfam; PF00089; trypsin_1.
DR      PRINTS; PR00722; CHYMOTRYPSIN.
DR      PRINTS; PR00018; KRINGLE.
DR      ProDom; PD000395; Kringle_1.
DR      SMART; SM00130; KR; 1.
DR      SMART; SM00020; Tryp_Spc; 1.
DR      PROSITE; PS00021; KRINGLE_1; 1.
DR      PROSITE; PS02400; KRINGLE_2; 1.
DR      PROSITE; PS50240; TRYPSIN_DOM; 1.
DR      PROSITE; PS00134; TRYPSIN_HIS; 1.

```

DR	PROSITE: PS00135; TRYPSIN_SER; 1	
KW	PLASMINOGEN activation; Hydrolyase; Serine protease; Glycoprotein;	
KW	Kringle; Signal; Multigene family.	
FT	SIGNAL	1
FT	CHAIN	37
FT	DOMAIN	45
FT	ACT SITE	142
FT	ACT SITE	189
FT	ACT SITE	238
FT	ACT SITE	345
FT	DISULFID	45
FT	DISULFID	126
FT	DISULFID	66
FT	DISULFID	97
FT	DISULFID	131
FT	DISULFID	174
FT	DISULFID	190
FT	DISULFID	182
FT	DISULFID	276
FT	DISULFID	308
FT	DISULFID	341
FT	CARBOHYD	315
SO	SEQUENCE	394 AA; 44105 MW; 9CCDD6F52F3D81FDC CRC64;

Query Match	31.4%;	Score 752;	DB 1;	Length 394;
Best Local Similarity	42.0%;	Pred. No. 2.9e-52;		
Matches 155;	Conservative 56;	Mismatches 134;	Indels 24;	Gaps 7;

[illegible]

Search completed: March 18, 2004, 11:16:59
Job time : 28.407 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:09:25 ; Search time 117.888 Seconds

(without alignments)
1153.535 Million cell updates/sec

Title: US-10-076-421-2

Perfect score: 2394
Sequence: 1 MALLARLLLCVIVSDSKG.....VSHFLPWRSHTEKENGAL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_rvirus:*
16: SP_bacteriophage:*
17: SP_archaeal:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1992	83.2	433	6	Q8MIL0	O8MIL0 oryctolagus
2	1987	83.0	433	6	Q8MIL7	Q8MIL7 oryctolagus
3	982.5	41.0	214	6	Q9X770	Q9X770 oryctolagus
4	882	36.8	231	11	Q8C6L2	Q8C6L2 mus musculus
5	871	36.4	154	4	Q968E8	Q968E8 homo sapien
6	870.5	36.4	516	4	Q9BU99	Q9BU99 homo sapien
7	868.5	36.3	562	4	Q86YK8	Q86YK8 homo sapien
8	826	34.5	564	6	Q8MK31	Q8MK31 oryctolagus
9	822.5	34.4	562	6	Q8SG23	Q8SG23 sus scrofa
10	784.5	32.8	395	4	Q9B2M1	Q9B2M1 sus scrofa
11	731.5	30.6	540	13	Q800Y7	Q800Y7 meleagris g
12	727	30.4	653	11	Q8VCS4	Q8VCS4 mus musculus
13	693.5	29.0	616	6	Q97507	Q97507 sus scrofa
14	689	28.8	615	4	Q81Z25	Q81Z25 homo sapien
15	680.5	28.4	597	11	Q35727	Q35727 mus musculus
16	680.5	28.4	609	11	Q80YCS	Q80YCS mus musculus

17	677	28.3	157	6	Q9TV48	Q9TV48 bos taurus
18	674.5	28.2	517	11	Q8K0D2	Q8K0D2 mus musculus
19	661.5	27.6	560	4	Q14520	Q14520 homo sapien
20	585	24.4	128	6	Q97587	Q97587 oryctolagus
21	498	20.8	810	4	Q15146	Q15146 homo sapien
22	496	20.7	454	6	Q46506	Q46506 papio hamad
23	496	20.7	812	11	Q9R0W3	Q9R0W3 rattus norv
24	495.5	20.7	300	4	Q96EP3	Q96EP3 homo sapien
25	495	20.7	103	6	Q9SM89	Q9SM89 equus caball
26	491.5	20.5	429	13	Q8AVB0	Q8AVB0 brachydanto
27	482.5	20.2	334	6	Q46507	Q46507 papio hamad
28	472.5	19.7	806	6	Q18783	Q18783 macropus eu
29	463.5	19.4	868	5	Q9YIV3	Q9YIV3 polyandroca
30	442.5	18.5	327	4	Q8N171	Q8N171 homo sapien
31	435.5	18.2	284	4	Q8NFR6	Q8NFR6 homo sapien
32	433.5	18.1	761	11	Q99J08	Q99J08 rattus norv
33	426.5	17.8	267	5	Q9BK47	Q9BK47 luidia foli
34	426	17.8	558	4	Q86YK4	Q86YK4 homo sapien
35	425.5	17.8	505	5	Q966V4	Q966V4 halocynthia
36	424	17.7	537	4	Q9BYE1	Q9BYE1 homo sapien
37	420	17.5	471	11	Q8CFE0	Q8CFE0 mus musculus
38	420	17.5	581	4	Q9BYE2	Q9BYE2 homo sapien
39	418.5	17.5	277	11	Q80WM7	Q80WM7 mus musculus
40	417.5	17.4	425	13	Q804X7	Q804X7 gallus gall
41	414	17.3	638	11	Q8R0P5	Q8R0P5 mus musculus
42	411.5	17.2	276	11	Q8CGR6	Q8CGR6 mus musculus
43	409	17.1	371	11	Q8CJ16	Q8CJ16 rattus norv
44	409	17.1	445	11	Q8CJ17	Q8CJ17 rattus norv
45	405.5	16.9	422	4	Q8WYCI	Q8WYCI homo sapien

ALIGNMENTS

RESULT 1			
Q8MIL0			
ID	Q8MIL0	PRELIMINARY;	PRT; 433 AA.
AC	Q8MIL0;		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Urokinase-type plasminogen activator.		
GN	PLAU.		
OS	Oryctolagus cuniculus (Rabbit).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
OX	NCBI_TaxId=9986;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22155945; PubMed=12149463;		
RA	Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,		
RA	Dichek D.A.;		
RT	"Increased expression of urokinase during atherosclerotic lesion		
RT	development causes arterial constriction and lumen loss, and		
RT	accelerates lesion growth."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.		
CC	-1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.		
EMBL	AY122285; AM83187.1; -		
DR	GO; GO:004263; F:chymotrypsin activity; IEA.		
DR	GO; GO:0016301; F:kinase activity; IEA.		
DR	GO; GO:0008293; F:peptidase activity; IEA.		
DR	GO; GO:004295; F:trypsin activity; IEA.		
DR	GO; GO:000608; P:proteolysis and peptidolysis; IEA.		
DR	InterPro: IPR009003; Cys_Ser_trypsin.		
DR	InterPro: IPR006209; EGF-like.		
DR	InterPro: IPR000001; Kringle.		
DR	InterPro: IPR001254; Peptidase S1.		
DR	InterPro: IPR001314; Peptidase S1A.		
DR	InterPro: IPR008293; Pept_S1A_UPA.		
DR	Pfam; PF00051; kringle; 1.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPSIN.		

PRINTS; PRO0018; KRINGLE.
 DR Prodom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS00022; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR PIRSF; PIRSF001144; Utk_Plaam_act; 1.
 DR Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
 KW SEQUENCE 433 AA; 48375 MW; 65E64F3641554990 CRC64;
 SQ

Query Match 83.2%; Score 1992; DB 6; Length 433;
 Best Local Similarity 83.4%; Pred. No. 1.6e-182;
 Matches 361; Conservative 26; Mismatches 44; Indels 2; Gaps 1;

1 MRALLARLLLCVLYVSDSKSGNELHOY--PSNCDCLNGGTCVSNKFFSNIHMNCNCPKFG 58
 1 MRVLLVCLLCALVVDSESGSHELHGVSDASNGCLNGGTCVYKFFSNIWRCNCPKFG 60

59 GQCEIDKSKTCYENGHFFRGKASTDTMGRPCLPNNSATVLCQTYHAHSDALQGLGK 118
 61 GEHCEIDLTCTCYHGDHSTRKANITIMRPLCLAMNSANVLTQTYHAHSDALQGLGK 120

119 HNYCRNPDNRPRWCYOVGLKPLVOECWHDCADEKSPSPPEELKFGCGQKTLRPRFK 178
 121 HNYCRNPDHQRPRWCYOVGLKPLVOECKVHDCSSGKKPALPCKLEFGCGQKTLRPRFK 180

179 IIGGEFTTIENOPFAIYRRHGGSVTYVCGSLISPCWVTSATHCFTIDYPKKEDYIV 238
 181 IIGGEFTTIENOPFAIYRRHGGSVTYVCGSLISPCWVTSATHCFTIDYPKKEDYIV 240

239 LGRSLNSNTOGEMKFEVENLILHKDYSADTLAHNDIALIKIRSKEGRCAPSRITQTI 298
 241 LGRSLNSMTPEGEMKFEVEOLILHGYRADTLAHNDIALIKIRSKEGRCAPSRITQTI 300

299 CLPSMTNDPQGTSCETITGFGKENDTYLPEQLKMTVVKLISHRCQCPHYGSEVTTK 358
 301 CLPSMTNDPQGTSCETITGFGKENDTYLPEQLKMTVVKLVSYOECQCPHYGSEVTTK 360

359 MLCADPQWKTDSCQGSGLPVCSLOGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPW 418
 361 MLCADPQWKTDSCQGSGLPVCSLOGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPW 420

419 IRSHTKEENGAL 431
 421 IRSHTKEENGAL 433

RESULT 2

Q8MHY7 PRELIMINARY; PRT; 433 AA.
 AC Q8MHY7;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Urokinase-type plasminogen activator.
 GN UROKINASE.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OC NCB1; Taxid=9986;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Yano W., Watanabe M.;
 RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
 complete cde.";

Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RL -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; AY029517; AK40239.1; -
 DR EMBL; AB087224; EAC02685.1; -
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR008293; Pept_S1A_uPA.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO00722; CHYMOTRYPSIN.
 DR PRINTS; PRO0018; KRINGLE.
 DR Prodom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR PIRSF; PIRSF001144; Utk_Plaam_act; 1.
 DR Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
 KW SEQUENCE 433 AA; 48444 MW; 6DD35A571010A6EE CRC64;
 SQ

Query Match 83.0%; Score 1987; DB 6; Length 433;
 Best Local Similarity 83.1%; Pred. No. 4.8e-182;
 Matches 360; Conservative 26; Mismatches 45; Indels 2; Gaps 1;

1 MRALLARLLLCVLYVSDSKSGNELHOY--PSNCDCLNGGTCVSNKFFSNIHMNCNCPKFG 58
 1 MRVLLVCLLCALVVDSESGSHELHGVSDASNGCLNGGTCVYKFFSNIWRCNCPKFG 60

59 GQCEIDKSKTCYENGHFFRGKASTDTMGRPCLPNNSATVLCQTYHAHSDALQGLGK 118
 61 GEHCEIDLTCTCYHGDHSTRKANITIMRPLCLAMNSANVLTQTYHAHSDALQGLGK 120

119 HNYCRNPDNRPRWCYOVGLKPLVOECWHDCADEKSPSPPEELKFGCGQKTLRPRFK 178
 121 HNYCRNPDHQRPRWCYOVGLKPLVOECKVHDCSSGKKPALPCKLEFGCGQKTLRPRFK 180

179 IIGGEFTTIENOPFAIYRRHGGSVTYVCGSLISPCWVTSATHCFTIDYPKKEDYIV 238
 181 IIGGEFTTIENOPFAIYRRHGGSVTYVCGSLISPCWVTSATHCFTIDYPKKEDYIV 240

239 LGRSLNSNTOGEMKFEVENLILHKDYSADTLAHNDIALIKIRSKEGRCAPSRITQTI 298
 241 LGRSLNSMTPEGEMKFEVEOLILHGYRADTLAHNDIALIKIRSKEGRCAPSRITQTI 300

299 CLPSMTNDPQGTSCETITGFGKENDTYLPEQLKMTVVKLISHRCQCPHYGSEVTTK 358
 301 CLPSMTNDPQGTSCETITGFGKENDTYLPEQLKMTVVKLVSYOECQCPHYGSEVTTK 360

359 MLCADPQWKTDSCQGSGLPVCSLOGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPW 418
 361 MLCADPQWKTDSCQGSGLPVCSLOGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPW 420

419 IRSHTKEENGAL 431
 421 IRSHTKEENGAL 433

RESULT 3

Q9XT70 PRELIMINARY; PRT; 214 AA.
 ID Q9XT70

```

AC 09XT70;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN NCBI_TaxID=9986;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin Y., Idell S.;
RT "Partial mRNA of rabbit uPA."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; A039351.1; -.
DR HSSP; P00749; 1EUN.
DR MEROPS; S01.231; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KM Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ
NON_TER 214
SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 41.0%; Score 982.5; DB 6; Length 214;
Best Local Similarity 82.8%; Pred. No. 5.3e-86;
Matches 178; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

QY 87 MGRPCLPMSATVLOQTYAHRSDALQLGKINYNCPNRRRPMWCYVGVGKPLVQEC 146
DB 1 MDRPCLAMNSANVLTQYAHRPDALQLGKINYNCPNRRRPMWCYVGVGKPLVQEC 60

QY 147 MNRDCAADGKRPSPPELAKQCCGKTLRPRFKIIGSEFTTIENQWPAALYRRHRGGSVT 206
DB 61 KVRD-SSGKKRPALEPGLRFGCCQKALRPRFKIIGSEFTTIENQWPAALYRRHRGGSVT 119

QY 207 YVCGSLISPCWYISATHCFIDYPKKEDYIVYGRSLNNTGQEMKFEVENTILHKDYS 266
DB 120 YVCGSLISPCWYISATHCFIDYPKKEDYIVYGRSLNNTGQEMKFEVENTILHKDYS 179

QY 267 ADTLAHHNDIALKIRSKRGCAQPSRTIQTICLP 301
DB 180 ADTLAHHNDIALKIRSKRGCAQPSRTIQTICLP 214

RESULT 4
Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC 08C6L2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium;
RA THE RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK054349; BAC35743.1; -.
DR PIR; P70534; P70534.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

Query Match 36.8%; Score 882; DB 11; Length 231;
Best Local Similarity 67.7%; Pred. No. 2.6e-76;
Matches 155; Conservative 27; Mismatches 45; Indels 2; Gaps 2;

QY 1 MRLLARLLCYLVSDSKGSENLHGV-PNCDLNGTCVSKNYSNIHWCNCPKKEGG 59
DB 1 MKWVLASLFLCALVNVNSEGSVLGAPDSESNCCQNGVCVSKYFSRIIRCSGCPKKEGG 60

QY 60 QHCEIDSKTCYGVGNHGFPRGKASTDTPMGRPCLPMSATVLOQTYAHRSDALQLGK 119
DB 61 EHCIEDASKTCYGVGNHGFPRGKASTDTPMGRPCLPMSATVLOQTYAHRSDALQLGK 120

QY 120 NYCRNPNDRRRPMWCYVGVGKPLVQECMVDCAADGKRPSPPELAKQCCGKTLRPRFKI 179
DB 121 NYCRNPNDRRRPMWCYVGVGKPLVQECMVDCAADGKRPSPPELAKQCCGKTLRPRFKI 180

QY 180 IGSEFTTIENQWPAALYRRHRGGS-VTVYCGSLISPCWYISATHCFI 227
DB 181 VGEFTEVENQWPAALYQKNKGSSPSFCGSLISPCWYISATHCFI 229

RESULT 5
Q96SE8 PRELIMINARY; PRT; 154 AA.
AC Q96SE8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Fu J., Bai X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human
RT urokinase-type plasminogen activator."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

Bai X., Fu J., Wang W., Xi X., Ruan C.;
 RT Overexpression of the amino-terminal fragment of human urokinase-type
 RT plasminogen activator in breast cancer cells results in decreased
 RT tumor invasion, growth and angiogenesis."
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL: AY029537; AAK38734.1; -
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR Pfam: PF00051; Kringle; 1.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00130; KR; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 KM Glycoprotein; Kinase; Kringle.
 SQ SEQUENCE 154 AA; 17305 MW; A3CF2FCFF505572 CRC64;

Query Match 36.4%; Score 871; DB 4; Length 154;
 Best Local Similarity 99.4%; Pred. No. 1.8e-75;
 Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRALLALLLCVLYVSDSKSGNELHGVPSNCDCLNGGTCVSNKYFSPNHNMCNPKKRGQ 60
 DB 1 MRALLALLLCVLYVSDSKSGNELHGVPSNCDCLNGGTCVSNKYFSPNHNMCNPKKRGQ 60
 QY 61 HCEIDSKTCYEGNGHFRGKASTDTMGKPCLPNNSATVLOQTTHARSDALQGLGKN 120
 DB 61 HCEIDSKTCYEGNGHFRGKASTDTMGKPCLPNNSATVLOQTTHARSDALQGLGKN 120
 QY 121 YCRNPDRRRPWCYVQGLKPLVQECWVHDCADG 154
 DB 121 YCRNPDRRRPWCYVQGLKPLVQECWVHDCADG 154

RESULT 6

Q9BU99 PRELIMINARY; PRT; 516 AA.
 AC Q9BU99;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=skin;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC EMBL: BC002795; AA02795.1; -
 DR HSSP; P00750; IASH.
 DR GO: GO:0004263; F:chymotrypsin activity; IEA.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00051; Kringle; 2.
 DR Pfam; PF00089; trypsin; 1.

DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 2.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00130; KR; 2.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00021; KRINGLE_1; 2.
 DR PROSITE: PS00070; KRINGLE_2; 2.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KM EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KM Serine protease.
 SQ SEQUENCE 516 AA; 57370 MW; BAB11901FDC96800 CRC64;

Query Match 36.4%; Score 870.5; DB 4; Length 516;
 Best Local Similarity 37.8%; Pred. No. 9.2e-75;
 Matches 195; Conservative 57; Mismatches 165; Indels 99; Gaps 12;

QY 2 RALLAALLLCVLYVSDSKSGNELH-----QVPSNCDCLNGGTCVSNKYFSPNHN 50
 DB 6 RGLCVLLLCGAVFVPSQ---EIHARFRGARASYGCSBPRCFNGCTQOALYFSDP-V 61
 QY 51 CNCPKFFGQHCEDSKTCYEGNGHFRGKASTDTMGKPCLPNNSATVLOQTTHARSD 110
 DB 62 CQCPBGFAGKCGCIDIRATCYEDQGI:SYRGWSTASGAECTWMSALAKRYSGRRPD 121
 QY 111 ALQGLGKNHYCNPDRRRPWCYVQGLKPLVQECWVHDCADG----- 154
 DB 122 AIRLGANNHYCNPDRRPSKPCVFRKAGYSSEFCSTPACSGNSDCYFGNSAYRGTH 181
 QY 155 -----KKPS-----SPPEELK----- 165
 DB 182 SLTESGASCLPNSMLIGKVTTAQNPSAOLGLGKNYCRANDGAKPRCHVLKNRRL 241
 QY 166 -----FOCG-OKTLRPRFKIIGEFITINOPFAIYRRH-RGGSVTVYCGSLI 214
 DB 242 WEYCDVPSCSTGGLRQYSGQFRIRKGLFADIAHPQAAIFAKHRRSPDERFLCGIILI 301
 QY 215 SPCWVISANTHCFIDYPKEDYIYVGRSLNSTQGEKREVENLILHKOYADTLAHNN 274
 DB 302 SSMWILSAHCFQERPPPHLTVILGRYRVVGESEQKEVRYVHKEFDDT--YDN 359
 QY 275 DIALTKRSKEGCAOPSRTIOTICLPWYNDRPOFGSCETIFGKENSTDYLYPEQLKM 334
 DB 360 DIALQKSDSSKCAQESSVRYVCLPPADLQLPDWTECHLSGKGNHALSPYSERLKE 419
 QY 335 TVVKLISHRECQOPHYVYGSVYTKMLCAAD-----PQWKT-DSQGDGSGPLVCSLGRM 388
 DB 420 AHVRLVPSRSCTGHLNLRVTIDNMLCAGDTRSGSPQANLHDACQSGSGPLVCLNDGRM 479
 QY 389 TLVGIVSWGRCALKDKPQVYTRVSHLPWIRSHTK 424
 DB 480 TLVGIIISWGLGCGQKDVPGVYTVVTLWIRDNMR 515

RESULT 7

Q86YK8 PRELIMINARY; PRT; 562 AA.
 AC Q86YK8;
 DT 01-JUN-2003 (TREMblrel. 24, Created)
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Tissue plasminogen activator.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Liu Y., Xu L., Zeng Y., He X.;

Db 252 TARQSNALALGKKNYCRNPDGSKPMCHVLKNNKLVLEYCDVQCA----- 299
 QY 165 KFOCG-QKTLRPRFKIIGSEFTTIENOPWPAIY-RRHRGGSVTVVCGSLISPCWVISA 222
 Db 300 --TCLLRDQKQPFKIKGLFTDITAHFWQAAIFNNRNSPQERLKGIIINSCWVISA 357
 QY 223 THCFIDYRKEDYIVYIGRSRLNSNTQGEKMEFEVENLIHKDYSADTLAHNDIALKIR 282
 Db 358 AHCFERFPQOKLRLVILGRVYPLVGAEEQIFVEVQPIIHERFDEGT--YNDIALALKLK 415
 QY 283 SKEGCAQPSRTIQTICLPSMTNDPQFGTSCGHTGFGKENSVDIYLPQQLKATVYKLSH 342
 Db 416 STSGCAQESQAVRLVCLPDLASLQIPDWTECELSGYGHEEFPSPVFSQLEAHVRLYPS 475
 QY 343 RECCQPHYGSEVTLTKMLCAADPCW-----KTDSGCGDSGGLVYCSLQGRMTLTGIYSW 396
 Db 476 SRCTFQQLKNNRTVIGNMCLAGDTRSGGAQVNLHDAQCDSGGLVCMQDMGHTLIGIISW 535
 QY 397 GRGCLAKDKPGVYTRVSHPLPWIRSH 422
 Db 536 GLGCGQKDVPGVYTRVNVYLGWIOQH 561

RESULT 9

Q98Q23 PRELIMINARY; PRT; 562 AA.
 AC Q98Q23;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE T-plasminogen activator.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxId=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain1 organ;
 RA Ding Y., Xue J., Bartlett J.D.;
 RL "T-plasminogen activator in tooth tissues."
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; AF364605; AAM00297.1; -.
 DR HSSP; P00761; IAN1.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_ser_clypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000083; Fibrinctn1.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00058; fn1; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_3; 1.
 DR PROSITE; PS01186; EGF_3; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KM EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KM Serine protease.
 SQ SEQUENCE 562 AA; 63668 MW; F9EBB4C77CB101B8 CRC64;

Query Match 34.4%; Score 822.5; DB 6; Length 562;
 Best Local Similarity 35.2%; Pred. 4.2e-70;
 Matches 173; Conservative 61; Mismatches 168; Indels 89; Gaps 8;

QY 20 GSNBLHQP-SNCD---CLNGGTCVSNKTFNSINHCNCPKKGQGHCEIDKSTCYBANG 75
 Db 74 GQTQCHSVPAKSCSBPCFNGJTCLOAIYFSD-VCQCPVGFGRQCEIDARATCYEDQG 132
 QY 76 HFRGKASTDTWRCPLPMSATVLTQTYAHNSDALDGLGKNNCRPNDRRRRCVCY 135
 Db 133 IYTRGTWSTTESGAECEVNNMTSGIASNPYNGRPDAVKGLGNHNYCRNPDKSKFWCY 192
 QY 136 QVGLKPLVQECVNHDCADGKKPSSPPEELKFQ----- 167
 Db 193 FKAERKSPDCSTPACTKKEECYTGKGLDYRGRSLTWSGARCLPMNSLVLMGKITYAM 252
 QY 168 -----CGQKTLR-PRFKII 180
 Db 253 NSNAQTLGLGKKNYCRNPDGDTQPMCHVLKDKRLTWYCDLPQCVTCGLRQYKEPQFRIK 312
 QY 181 GGEFTTIENOPWPAIY-RRHRGGSVTVVCGSLISPCWVISA THCFIDYPKKEDYIVYL 239
 Db 313 GGLYADITSPWQAAIFVKNRSPGERFLCGGIISSCVLSAHHFQGRPPHHVRLV 372
 QY 240 GRSRLNSNTQGEKFEVENLIHKDYSADTLAHNDIALIKIRSKGRCQAQPSRTIQTIC 299
 Db 373 GRTRYLVPGHEEQAFVEKTYVHKEPDDT--YNDIALLOKSDSLTCAQESDAVRTVC 430
 QY 300 LPSMTNDPQFGTSCGHTGFGKENSVDIYLPQQLKATVYKLSHRECOQHYYGSVTTTCM 359
 Db 431 LPEANILQIPDWTECELSGYGHEEFPSPVFSERKEAHVRLYBSSRTSKHLFKKTIYNNM 490
 QY 360 ICAADPCW-----KTDSGCGDSGGLVYCSLQGRMTLTGIVSGRCALDKRGVYTRVS 413
 Db 491 LCGADTRSGGDNANLHDAQCDSGGLVCMKGNHMTLVGVISKGLCGCGQKDVPGVYTKYT 550
 QY 414 HPLPWIRSHTK 424
 Db 551 NYLMTWRDNT 561

RESULT 10
 Q9BZM1 PRELIMINARY; PRT; 395 AA.
 AC Q9BZM1;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Neonatal thrombolytic agent alpha-form (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dou D.;
 RL "A brain-type plasminogen activator."
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; AF260825; AAK11956.1; -.
 DR HSSP; P00750; IPR2.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR000083; Fibrinctn1.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00039; fn1.1.
 DR Pfam: PF00051; kringle.1.
 DR Pfam: PF00089; trypsin.1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle.1.
 DR SMART: SM00058; FN1.1.
 DR SMART: SM00130; KR.1.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS01253; FIBRONECTIN_1.1.
 DR PROSITE: PS00021; KRINGLE_1.1.
 DR PROSITE: PS00021; KRINGLE_2.1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR GlycoProtein: Hydrolase; Kringle; Protease; Serine protease.
 FT NON_TER 395
 SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 32.8%; Score 784.5; DB 4; Length 395;
 Best Local Similarity 42.0%; Pred. No. 1.2e-66;
 Matches 161; Conservative 49; Mismatches 144; Indels 29; Gaps 8;

QY 50 WCNCPPKFGGCHGHEIDSKTCYVNGHFFRGKASTDTMGRPCLPMSATVLOQTTHAHS 109
 DB 33 WCNS---GAQGS--EGNSDCYFGNSAYVGTSLTESGASCLPMSMILIGVYTAQNP 87
 QY 110 DALQLGKHNKYNCRNDNRNRRPWCYVQVGLKPLVQECMVNDGDKKPSPRELKFOCG 169
 DB 88 SAQALGKHNKYNCRNDNRNRRPWCYVQVGLKPLVQECMVNDGDKKPSPRELKFOCG 169
 QY 170 -QKTLRPFKIIGSEFTTLENQWPAIYRRH-RGSSVTVVCGSLISPCWISATHCFI 227
 DB 134 LRQYQPOFQIRKIGLFLADISHPMQAAIFAKHEESPCEPRLCGIILISCMILSAHCFQ 193
 QY 228 DYKEDDIYVILGRSLNSTQGEKFEVENLLHNDYADTLAHNDILALKIRSEGR 287
 DB 194 ERPPPHLLVILGRYRVVVGEEBQKFECEKTYLVHKEFDDT--YNDIALLLQKSDSSR 251
 QY 288 CAQPSRTIQTCLPSMYNDPQFGTSCETGFGKENSIDYLYPQKMTVVKLISHBECQ 347
 DB 252 CAQSSVSRVCLPPADLPQPMTEGCLSGYGHKALSPYSERLKAHRLYPSRCTS 311
 QY 348 PHYGSEVTTKMLCAAD-----PQWKT-DSQGDSSGGLVCSLOGRWTLTGIVSWGRC 401
 DB 312 QHLNRTVTDNMLCAGDTRSGQFQAMLHACQDSSGGLVCLNDGRWTLVGIIISWGLGCG 371
 QY 402 LKDKRGVYTRVSHFLPMIRSHTK 424
 DB 372 QKQVPGVYTKVTNYLDMIRDMR 394

RESULT 11

Q800Y7 PRELIMINARY; PRT; 540 AA.
 AC Q800Y7;
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Hepatocyte growth factor activator (Fragment).
 OS Meleagris gallopavo (Common turkey).
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.
 OC NCBI_TaxId=9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21212796; PubMed=12128063;
 RA Holtsberger D.R., Becker A.E., Thureston R.J., Rice C.D.;

RT "Expression of a hepatocyte growth-factor activator protein in turkey
 (Meleagris gallopavo) deferent duct epithelial cells."
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RT Comp. Biochem. Physiol. 132:769-777 (2002).
 RN [2]

RP SEQUENCE FROM N.A.

RA Holtsberger D.R., Becker A.E., Thureston R.J., Rice C.D.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY216598; AA046038.1;
 DR GO: GO:0005576; C:extracellular; IEA.
 DR GO: GO:0004263; F:chymotrypsin activity; IEA.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR00742; EGF_2.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR000083; Fibrinctn1.
 DR InterPro: IPR000562; FN_Type_II.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00051; kringle.1.
 DR Pfam: PF00089; trypsin.1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle.1.
 DR SMART: SM00058; FN1.1.
 DR SMART: SM00059; FN2.1.
 DR SMART: SM00130; KR.1.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01253; FIBRONECTIN_1.1.
 DR PROSITE: PS00021; KRINGLE_1.1.
 DR PROSITE: PS00070; KRINGLE_2.1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 FT NON_TER 1
 FT NON_TER 540
 SQ SEQUENCE 540 AA; 61159 MW; 0BB3BA489C0B577F CRC64;

Query Match 30.6%; Score 731.5; DB 13; Length 540;
 Best Local Similarity 37.2%; Pred. No. 2.2e-61;
 Matches 159; Conservative 67; Mismatches 166; Indels 35; Gaps 11;

QY 11 CVLVSDSKGSNELHQVPSNCDCLNGTCTC-----VSNKYPSNIHMCNCPKPKFGGCHGEIDK 66
 DB 121 CTCTVDQGMCLAVHEHKSCHDPCNMGCECKMIASSGKT-----CDCKGPFVGKYNIV 175
 QY 67 SKTCYBNGHFFRGKASTDTMGRPCLPMSATVLOQTTHAHSDALQLGKHNKYNCRND 126
 DB 176 NHHCYRNGTEYRGTAKTITISGHSCLPMSDLYRLHVDSEKAAQVLGIGPSSYCRND 235
 QY 127 NRRRPWCYVQVGLKPLVQECMVNDGDKKPSPREL-----KFGCGCK-----TLRP 175
 DB 236 EDEKPMCTYIMKONSLSWEYCNITSCASRRRPPVLEDTTFPAVRPCCGRHKKGSFVRP 295
 QY 176 RFKIIGSEFTTLENQWPAIYRRHSGSVTVVCGSLISPCWISATHCFIDYPKEDY 235
 DB 296 R--IIIGSSSLGSHHWTAIY---IGES---FCAGTLIQTCTWVSAHCFANSPKSSSI 347
 QY 236 IYLLGRSLNSTQGEKFEVENLLHNDYADTLAHNDILALKIRSEGRCAQSSRTI 295
 DB 348 KYVLGHPNRTYDVTQTEIRKYLIPQYSVFRPTEN-DIALIKLKXNGQRCAYVSGQEV 406
 QY 296 QITCLP---SMYNDPQFGTSCETGFGKENSIDYLYPQKMTVVKLISHBECQPHYHG 352
 DB 407 QPCLPESNTVPD-QF--KQISGKHGHNITGSDVLOQTTLITIEBKCRSPETIG 463
 QY 353 SEVTKMLCAADPQWKTDSQGDSSGGLVCSLOGRWTLTGIVSWGRCALKDKPGVYTRV 412

```
Db 464 TEISENMFCCGYRDPDSKSDACGDSGGPLACENNEISITLYGVISWGDCGKVNKPGYTRY 523
Qy 413 SHFLPMI 419
Db 524 PNYVNMV 530

RESULT 12
Q8VCS4 PRELIMINARY; PRT; 653 AA.
ID Q8VCS4
AC Q8VCS4;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strauberg R.;
CC Submitted (DSC-2001) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1 SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSSP; P00761; IAN1.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibronctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTYPET1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FRI8D90174ED6FDD CRC64;
```

```
Qy 11 CVLVVSDSKSNELHVPENSCDLCNGTCVSNKFFSNIHW-----CNCPEKTGGQHC 62
Db 225 CGMEGQARCEDTHTHTACLSSPCLNGTC-----HLVGTGTSVTCGLGAGRC 275
Qy 63 EIDSKRTCYEGNGHFYRGA51DTWGRPCLPMNSATVLQOTYHAHSDALDLGKKNYC 122
Db 276 NIVPTEHCLIGNOTEXRGA5TAAGSLCLANMSIDLVOELHVD5VAALVLLGLGHAAC 335
Qy 123 RNPDRRRPWCYVQVGLKVLVEGCMWHDAC-GKKRSPPEEL-----KFCGQOK 171
Db 336 RNDKDERPCWCVVKNALSWECRLTACE5IARVH5G5PELIALPESAPRAVPCGGR 395
Qy 172 -----TLRPRFKLIGEFPTTIEQWFAIYRHRG5TYVYCGSLISPCWISATNCF 226
Db 396 HKKRTLRER--IIGSSSLPGSHPLAIIY---IGNS---FCAGSLVHTCWWVA5AHC 447
Qy 227 IDVPKEDYIYVIGSRINSNTQGENKFVEVNLILHKDYSADTLAHNDIALIKTRSKG 286
Db 448 ANSPRDSITVLGQHFNRRTDVTQTFGIEKVPYTLVSVPNNH-DLVILRLKKGE 506
Qy 287 RCAQPSRTIQTCLPSMNDPQRTSCETLPGKENSITDYLVEQLKMTVVKLISHREG 346
Db 507 RCAVR5QFVQPLCLPAGSSPFTGRCQIAGWMDENV5SYNSLLEALVPLVADHK5 566
Qy 347 QPHYSEVYTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSMGRGALKDRP 406
Db 567 SPFVYADISPNLCLGCFYDCKSDACGDSGGPLVEKNGAVLYVGI5WGDCGRLNKP 626
Qy 407 GYTVRVSHFLPWIRSHTK 424
Db 627 GYTVRVANVDMINDRIR 644

RESULT 13
ID 097507 PRELIMINARY; PRT; 616 AA.
AC 097507;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Takahashi T., Kihara T.;
RT "Porcine liver factor XII."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1 SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibronctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
```

DR Pfam: PF00040; fn2; 1.
 DR Pfam: PF00051; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00013; ENTPEPT.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; FN_Type_II; 1.
 DR ProDom: PD000395; kringle; 1.
 DR SMART: SM00181; EGF; 2.
 DR SMART: SM00058; FN1; 1.
 DR SMART: SM00059; FN2; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01253; FIBRONECTIN_1; 1.
 DR PROSITE: PS00023; FIBRONECTIN_2; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KM EGF-like domain; Glycoprotein; Hydrolyase; Kringle; Protease;
 Serine protease.
 SQ SEQUENCE 616 AA; 66012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 29.0%; Score 693.5; DB 6; Length 616;
 Best Local Similarity 35.3%; Pred. No. 1,1e-57;
 Matches 161; Conservative 61; Mismatches 159; Indels 75; Gaps 14;

QY QVPSNCDCLNGGTGVSNKYSNIHMCNCPKFGGHCIDSKTYEGNGHFRYKASTD 85
 DB QVCTSTPCLNGLSGSLOTE---GHRLCGCPGVAGRLCDVDLKERCVSDRGSLRGMAT 232
 QY 86 TMRPCLPMSATVLOQTY-HARRSALQGLGKHYCNRPDNRKRWCVYVGLKFLVQ 144
 DB 233 LSGAPCOPMAS---EATYNNMTAEQALNWLGLDHAFCRNPDDTRPWCFCVWRGDDLSWQ 288
 QY 145 ECVVHPCAD--GKKP-----SSPEEL- 164
 DB 289 YCLARCOAIGEPPLITLTQSBSEKQDPLSRBPQTTPPSQNLTKMCAKPPQRG 348
 QY 165 -----KFGCGKTLRPRF---KIIGSEFTTIENQWFAIYRRHGGSVTVYCGSLI 214
 DB 349 PLPSAGLVGGGR-LRRRLSLNRIVGGLVALPGAHYIALYWGQV-----FCAGSLI 401
 QY 215 SPQWVSATHCFTDYPKEDYIYLGSRNLNSTOGEMKEVENLIHKQYADTLAHN 274
 DB 402 APCWVLAHAHCLNRPAPBELTVVLGQDRHNSCEQCOTLAIVRSYRLHESYSPTYQH-- 459
 QY 275 DIALLKIR-SKGRCAQPSRTIOTICP---SMYNDQPGTSCETIGFGKENSIDVLYPE 330
 DB 460 DIALVRKETADGCAHSPFPVQCLPSVASSAPE-GALCEVAMGWFQFGAEYS 518
 QY 331 QLMTVTKLISHRECOQPHYGSEVTTKMLCAADPWKTKSCQDSSGGPLVC---SLQGR 387
 DB 519 FLDEAGVPLSPRSCADYHGAFTPGMLCAGTLEGTDAACGDSGGLVDCDETARQ 576
 QY 388 MTLTGIVSGRGALCKDKGVYTRVSHFLPWRSH 423
 DB 579 LVLRGIVSWGCGDRLKPGVYTDVANYLAMIQHT 614

RESULT 14

ID 081225 PRELIMINARY; PRT; 615 AA.
 AC 081225;
 DT 01-MAR-2003 (TREMUREL. 23, Created)
 DT 01-MAR-2003 (TREMUREL. 23, Last sequence update)
 DT 01-OCT-2003 (TREMUREL. 25, Last annotation update)
 DE Coagulation factor XII-Me.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wada H., Nishio K., Nakatani K., Kasei Y., Abe Y., Nobori T.,
 RT "Molecular characterization of coagulation factor XII-Me."
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB095845; BAC23095.1; -
 DR GO: GO:0005576; C:extracellular; IEA.
 DR GO: GO:0005576; C:calcium ion binding; IEA.
 DR GO: GO:0004293; F:chymotrypsin activity; IEA.
 DR GO: GO:0004293; F:trypsin activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR000083; Fibnctnl.
 DR InterPro: IPR000562; FN_Type_II.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00008; EGF; 2.
 DR Pfam: PF00039; FN1; 1.
 DR Pfam: PF00039; FN2; 1.
 DR Pfam: PF00040; FN2; 1.
 DR Pfam: PF00051; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00013; ENTPEPT.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; FN_Type_II; 1.
 DR ProDom: PD000395; kringle; 1.
 DR SMART: SM00181; EGF; 2.
 DR SMART: SM00179; EGF_CA; 2.
 DR SMART: SM00058; FN1; 1.
 DR SMART: SM00059; FN2; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01253; FIBRONECTIN_1; 1.
 DR PROSITE: PS00023; FIBRONECTIN_2; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 28.3%; Score 689; DB 4; Length 615;
 Best Local Similarity 34.3%; Pred. No. 3,1e-57;
 Matches 153; Conservative 59; Mismatches 164; Indels 70; Gaps 9;

QY 33 CLNGTGVSNKYSNIHMCNCPKFGGHCIDSKTYEGNGHFRYKASTDTMRPCL 92
 DB 183 CHHGRCLE---VEGRLCHCPVGTGPPCDVTKASCYDGRGLSYGLARTLLSGAPCQ 239
 QY 93 PMSATVLOQTY-HARRSALQGLGKHYCNRPDNRKRWCVYVGLKFLVQECVHDC 151
 DB 240 PMAS---EATYNNMTAEQARWNLGHAFCRNPDDTRPWCFLNDRLSWEYCOLAQ 295
 QY 152 -----ADGKKPSPEEL- 164
 DB 296 QPTQAAPPPIVPSRLHVLMPAQPAPKPPQPTTRTPPOSQTGALPAKREQPSLIRNG 355
 QY 165 KFGCGQ---KTLRPRKIIGSEFTTIENQWFAIYRRHGGSVTVYCGSLISPQWIS 221
 DB 356 PLSCGGRLKRLSLSMWRVVGGLVALRGAPHYIALYWGHS-----FCAGSLIAPCVLIT 409
 QY 222 ATTCFTDYPKEDYIYLGSRNLNSTOGEMKEVENLIHKQYADTLAHNDILAKI 281
 DB 410 AAHCLDRAPEDLTVLGOERRNHSCEQCOTLAIVRSYRLHFAFS--PVSQYDHLALRL 467

```

QY 282 R-SKEGRCAQPSRTIQTICPSMYNDPQFTSCETITFGKFNSTDYLYPEQLKMTVVKLI 340
DB 468 QEDADSGCALLSPVQVCLPSCAARSEITLQVAGCGHFGEGAEYSAFLQEAQVPL 527
QY 341 SHRECCQPHYYSSEVTTKMLCAADPQWKTDSCQDSSGSPVLSIQ--RWTLTGIVSWG 397
DB 528 SLRCSAPDVHSGSILPGMLCAGFLGEGTDACQDSSGSPVCEQAERRLTLQGIISWG 587
QY 398 RGCALKDKPGVYTRVSHFLPWIRSH 423
DB 588 SGGCDNRKPGVYTDVAYVYLAIRREHT 613

RESULT 15
ID 035727 PRELIMINARY; PRT; 597 AA.
AC 035727;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Schloesser M., Schwager S., Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X9571; CA67891.1; -.
DR HSSP; P00760; 1A07.
DR MEROPS; S01.211; -.
DR MGD; MGI:1891012; F12.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibinectin.
DR InterPro; IPR000562; FN_type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00068; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPR11.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000995; FN_type_II; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.

```

```

DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine Protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match 28.4%; Score 680.5; DB 11; Length 597;
Best Local Similarity 34.7%; Pred. No. 1,9e-56;
Matches 156; Conservative 65; Mismatches 169; Indels 59; Gaps 12;

QY 17 DSKGSNEL-----QVPCNCCDCLNGTCVSNKYFNSIMHCNCPKKFGGQHCEIDKSKTC 70
DB 162 ECKGS-EAHCKPVASQASINPCDNGSCL--LVEDHPLCRCPGTGTGFCDDIDMAIC 217
QY 71 YEGNGHFYRGKASTDTMGRPCLPMNSATVLAQTY-HAHSDALQGLGKANTCRNPDR 129
DB 218 YERGLSTYRGAGTGTSGAPCQRM--TVEATYRMTEKQALSMGLGHHAFCRNDPDT 273
QY 130 RFWCYVQVGLKPLVQECMVHDC-----ADGKRPSSRP----- 161
DB 274 RFWCFWMSGDRLSWDYCGLEOCPTTFAPLVVPESQESPSQAPSLSHAPNDSTDHQSL 333
QY 162 -EELKFGCGQ---KTLRPRFKIIGEFITIBNQPFALIRRHGGSVTVYVCGSLISPC 217
DB 334 SKTNMGCCQRFERKGLSSFMRVVGGVLALPGSHPIYALYKNN-----FCAGSLIAC 387
QY 218 WVISATHCFIDYPKEDYIVYIGRSRLNSNTQEMKEFEVENTILHKDYADTLAHHNDIA 277
DB 388 WVTIAHCLQNRAPAPBELTVVIGQDRHNSCEMCCOTLAVRSYALHGFSSITYQH--DLA 445
QY 278 LKIR-SKEGRCAQPSRTIQTICLPSMYNDPQFTSCETITFGKFNSTDYLYPEQLKMTV 336
DB 446 LRLQESKTNSCALISPHVQVCLPSCAARPSRTVCEVAGWGHOLEGAEEYSTFLQEAQ 505
QY 337 VKLISHREGQPHYYSSEVTTKMLCAADPQWKTDSCQDSSGSPVLSIQ--RWTLTGIVSG 393
DB 506 VPFIALDRGNSNVHGDALIPGMLCAGFLGEGTDACQDSSGSPVLCBEGTAHQLTRGV 565
QY 394 VSWRGCAKDKPGVYTRVSHFLPWIRSH 422
DB 566 ISWGSQGDNRKPGVYTDVAYVYLAIRREHT 613

```

Search completed: March 18, 2004, 11:19:17
 Job time: 119.888 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:53:24 ; Search time 1.65427 Seconds
(without alignments)
683.197 Million cell updates/sec

Title: US-10-076-421-3

Perfect score: 22

Sequence: 1 KKG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	4	5	AAg79461
2	22	100.0	7	4	AAm44844
3	22	100.0	7	4	AAm44233
4	22	100.0	7	4	AAm44849
5	22	100.0	7	4	AAm43769
6	22	100.0	7	4	AAm45363
7	22	100.0	7	4	AAm45898
8	22	100.0	7	4	AAm44584
9	22	100.0	7	4	AAm45903
10	22	100.0	7	4	AAm44641
11	22	100.0	8	4	AAm35204
12	22	100.0	10	4	AAm48320
13	22	100.0	10	4	AAm48341
14	22	100.0	10	4	AAm48341
15	22	100.0	10	4	AAm48341
16	22	100.0	10	4	AAm48341
17	22	100.0	11	2	AAm16616
18	22	100.0	12	2	AAm30313
19	22	100.0	12	5	AAO22420
20	22	100.0	13	2	AAm30274
21	22	100.0	13	2	AAm30274
22	22	100.0	13	5	AAE22487
23	22	100.0	14	2	AAm31742
24	22	100.0	14	6	AAm31742
25	22	100.0	15	2	AAm31741

26	22	100.0	15	2	AAm6300
27	22	100.0	15	2	AAy22245
28	22	100.0	16	2	AAm98109
29	22	100.0	17	2	AAm99102
30	22	100.0	17	2	AAm22589
31	22	100.0	17	6	AAm81913
32	22	100.0	18	2	AAm10659
33	22	100.0	18	2	AAm10661
34	22	100.0	18	2	AAm13929
35	22	100.0	18	2	AAm15796
36	22	100.0	18	2	AAm15798
37	22	100.0	18	2	AAm99101
38	22	100.0	18	2	AAm94000
39	22	100.0	18	2	AAm66317
40	22	100.0	18	2	AAm66319
41	22	100.0	18	2	AAy22250
42	22	100.0	18	2	AAy22248
43	22	100.0	19	2	AAm99100
44	22	100.0	19	2	AAm93999
45	22	100.0	19	3	AAy67472

ALIGNMENTS

RESULT 1
AAg79461
ID AAG79461 standard; peptide; 4 AA.
XX
AC AAG79461;
XX
DT 15-NOV-2002 (first entry)
XX
DE sc-uPA long A chain peptide fragment 1, amino acids 55-58.
XX
XX Single chain prepro-urokinase; sc-uPA; pro-urokinase; HIV;
XX high molecular weight urokinase-type plasminogen activator; HMW-uPA;
XX long A; long B; EGF-like domain; kringle domain; urokinase receptor;
XX low molecular weight urokinase-type plasminogen activator; LMW-uPA; CD87;
XX binding domain.
XX Homo sapiens.
XX OS
XX EP1232755-A2.
XX PN
XX 21-AUG-2002.
XX PD
XX 15-FEB-2002; 2002EP-00003555.
XX PF
XX 20-FEB-2001; 2001JP-00042655.
XX PR 19-JUN-2001; 2001JP-00184284.
XX PT
XX (JCRP-) JCR PHARM CO LTD.
XX PI
XX Wada M, Wada N;
XX WPI; 2002-610512/66.
XX DR
XX Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g. high
XX molecular weight urokinase-type plasminogen activator, amino-terminal
XX fragment or an anti-CD87 antibody.
XX PS
XX Example; Page 23; 38pp; English.
XX The sequences given in AAG79461-63 represent peptide fragments derived
XX from the long A chain of single chain prepro-urokinase (sc-uPA). Pro-
XX urokinase (amino acids 21-431) with a cleavage between amino acids 178
XX and 179 gives high molecular weight urokinase-type plasminogen activator
XX (HMW-uPA). HMW-uPA is a protein consisting of two peptide chains linked
XX by a di-sulphide bond. The chains, long A and B, are formed by enzymatic
XX cleavage between amino acids 178 and 179 of pro-urokinase. HMW-uPA
XX includes an EGF-like domain, a kringle domain and a urokinase receptor
XX (CD87) binding domain. HMW-uPA is then cleaved between amino acids 155

CC and 156 to give low molecular weight urokinase-type plasminogen activator
 CC (LMW-uPA) (amino acids 156-178 and 179-431), that has no plasminogen
 CC activator activity. SC-uPA, or these fragments of it, may be used in the
 CC anti-HIV agents of the invention which comprise a ligand molecule that
 CC binds to CD87. The agents are useful for treating HIV-infected humans for
 CC suppression of reproduction of HIV. The anti-HIV agents act by a
 CC mechanism of action different from those of conventional drugs, widening
 CC the choice of therapeutics agents and avoiding problems of resistant HIV
 CC XX

SO Sequence 4 AA:

Query Match 100.0%; Score 22; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
 DB 1 KKEG 4

RESULT 2

AA44844 standard; peptide; 7 AA.

AA44844; 25-OCT-2001 (first entry)

H11 binding site consensus conforming peptide (CCP) #1115.

XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
 KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
 KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
 KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
 KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.

XX Homo sapiens.
 OS Synthetic.

XX CA2290722-A1.

XX 08-JUN-2001.

XX 08-DEC-1999; 99CA-02290722.

XX 08-DEC-1999; 99CA-02290722.

XX (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
 PI Entwistle JM, Macdonald GC;

XX WPI; 2001-425937/46.

XX Composition useful for treating and diagnosing cancer, comprises stress
 PT protein-peptide complexes associated with tumor, and isolated antigen-
 PT binding fragments of an antibody that binds specifically to the complex.

XX Example 4; Page 105; 154pp; English.

XX The present invention describes a composition (I) comprising stress
 CC protein-peptide complexes (SPPC) associated with tumours that is
 CC specifically immunogenically cross-reactive with cell surface-associated
 CC SPPCs specific to target cancer (TC). Also described is an isolated
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs
 CC or a population of different SPPCs consisting of immunogenic cancer cell
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
 CC used in vaccine production and as a tumour-specific immunogenic response
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting

CC or imaging cancer cells, and to monitor the course of amelioration of
 CC malignancy in an individual. AA43707 to AA47109 represent peptides
 CC which are used in the exemplification of the present invention

SO Sequence 7 AA:

Query Match 100.0%; Score 22; DB 4; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
 DB 2 KKEG 5

RESULT 3

AA44233 standard; peptide; 7 AA.

AA44233; 25-OCT-2001 (first entry)

H11 binding site consensus conforming peptide (CCP) #504.

XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
 KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
 KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
 KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
 KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.

XX Homo sapiens.
 OS Synthetic.

XX CA2290722-A1.

XX 08-JUN-2001.

XX 08-DEC-1999; 99CA-02290722.

XX 08-DEC-1999; 99CA-02290722.

XX (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
 PI Entwistle JM, Macdonald GC;

XX WPI; 2001-425937/46.

XX Composition useful for treating and diagnosing cancer, comprises stress
 PT protein-peptide complexes associated with tumor, and isolated antigen-
 PT binding fragments of an antibody that binds specifically to the complex.

XX Example 4; Page 102; 154pp; English.

XX The present invention describes a composition (I) comprising stress
 CC protein-peptide complexes (SPPC) associated with tumours that is
 CC specifically immunogenically cross-reactive with cell surface-associated
 CC SPPCs specific to target cancer (TC). Also described is an isolated
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs
 CC or a population of different SPPCs consisting of immunogenic cancer cell
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
 CC used in vaccine production and as a tumour-specific immunogenic response
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
 CC or imaging cancer cells, and to monitor the course of amelioration of
 CC malignancy in an individual. AA43707 to AA47109 represent peptides
 CC which are used in the exemplification of the present invention

SO Sequence 7 AA:

Query Match 100.0%; Score 22; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRF 4
DB 2 KRF 5

RESULT 4

AAM4849
ID AAM4849 standard; peptide; 7 AA.

AC AAM4849;

DT 25-OCT-2001 (first entry)

DE H11 binding site consensus conforming peptide (CCP) #1120.

XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;

KW immunogenically cross-reactive; cancer; immunogenic cancer cell;

KW cytostatic; vaccine; tumour-specific immunogenic response inducer;

KW astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;

KW endymoma; medulloblastoma; primitive neural ectodermal tumour.

XX Homo sapiens.

OS Synthetic.

PN CA2290722-A1.

PD 08-JUN-2001.

PF 08-DEC-1999; 99CA-02290722.

PR 08-DEC-1999; 99CA-02290722.

PI (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

PI Entwistle JM, Macdonald GC;

XX WPI; 2001-425937/46.

XX Composition useful for treating and diagnosing cancer, comprises stress

PT protein-peptide complexes associated with tumor, and isolated antigen-

PT binding fragments of an antibody that binds specifically to the complex.

XX Example 4; Page 105; 154pp; English.

XX The present invention describes a composition (I) comprising stress

CC protein-peptide complexes (SPPC) associated with tumors that is

CC specifically immunogenically cross-reactive with cell surface-associated

CC SPPCs specific to target cancer (TC). Also described is an isolated

CC antigen-binding fragment of an antibody that binds specifically to SPPCs

CC or a population of different SPPCs consisting of immunogenic cancer cell

CC surface-associated SPPC of TC. (I) has cytostatic activity and can be

CC used in vaccine production and as a tumour-specific immunogenic response

CC inducer. (I) is useful for treating 71 types of cancers or tumors in a

CC subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,

CC oligodendroglioma, endymoma, medulloblastoma, and primitive neural

CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including

CC or imaging cancer cells, and to monitor the course of amelioration of

CC malignancy in an individual. AAM43707 to AAM47109 represent peptides

CC which are used in the exemplification of the present invention

XX Sequence 7 AA;

XX Query Match 100.0%; Score 22; DB 4; Length 7;

XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;

XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRF 4
DB 2 KRF 5

RESULT 5
AAM43769
ID AAM43769 standard; peptide; 7 AA.

AC AAM43769;

DT 25-OCT-2001 (first entry)

DE H11 binding site consensus conforming peptide (CCP) #40.

XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;

KW immunogenically cross-reactive; cancer; immunogenic cancer cell;

KW cytostatic; vaccine; tumour-specific immunogenic response inducer;

KW astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;

KW endymoma; medulloblastoma; primitive neural ectodermal tumour.

XX Homo sapiens.

OS Synthetic.

PN CA2290722-A1.

PD 08-JUN-2001.

PF 08-DEC-1999; 99CA-02290722.

PR 08-DEC-1999; 99CA-02290722.

PI (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

PI Entwistle JM, Macdonald GC;

XX WPI; 2001-425937/46.

XX Composition useful for treating and diagnosing cancer, comprises stress

PT protein-peptide complexes associated with tumor, and isolated antigen-

PT binding fragments of an antibody that binds specifically to the complex.

XX Example 4; Page 101; 154pp; English.

XX The present invention describes a composition (I) comprising stress

CC protein-peptide complexes (SPPC) associated with tumors that is

CC specifically immunogenically cross-reactive with cell surface-associated

CC SPPCs specific to target cancer (TC). Also described is an isolated

CC antigen-binding fragment of an antibody that binds specifically to SPPCs

CC or a population of different SPPCs consisting of immunogenic cancer cell

CC surface-associated SPPC of TC. (I) has cytostatic activity and can be

CC used in vaccine production and as a tumour-specific immunogenic response

CC inducer. (I) is useful for treating 71 types of cancers or tumors in a

CC subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,

CC oligodendroglioma, endymoma, medulloblastoma, and primitive neural

CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including

CC or imaging cancer cells, and to monitor the course of amelioration of

CC malignancy in an individual. AAM43707 to AAM47109 represent peptides

CC which are used in the exemplification of the present invention

XX Sequence 7 AA;

XX Query Match 100.0%; Score 22; DB 4; Length 7;

XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;

XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 6
ID AAM45363 standard; peptide; 7 AA.
XX
XX AAM45363;
XX
XX 25-OCT-2001 (first entry)
XX
XX H11 binding site consensus conforming peptide (CCP) #1634.
XX
XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
XX immunogenically cross-reactive; cancer; immunogenic cancer cell;
XX cytostatic; vaccine; tumour-specific immunogenic response inducer;
XX astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
XX ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX CA2290722-A1.
XX
XX 08-JUN-2001.
XX
XX 08-DEC-1999; 99CA-02290722.
XX
XX 08-DEC-1999; 99CA-02290722.
XX
XX 08-DEC-1999; 99CA-02290722.
XX
XX (NOVO-) NOVOPHARM BIOTECH INC.
XX
XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
XX Entwistle JM, Macdonald GC;
XX
XX WPI; 2001-425937/46.
XX
XX Composition useful for treating and diagnosing cancer; comprises stress
XX protein-peptide complexes associated with tumor, and isolated antigen-
XX binding fragments of an antibody that binds specifically to the complex.
XX
XX Example 4; Page 106; 154pp; English.
XX
XX The present invention describes a composition (I) comprising stress
XX protein-peptide complexes (SPPC) associated with tumours that is
XX specifically immunogenically cross-reactive with cell surface-associated
XX SPPCs specific to target cancer (TC). Also described is an isolated
XX antigen-binding fragment of an antibody that binds specifically to SPPCs
XX or a population of different SPPCs consisting of immunogenic cancer cell
XX surface-associated SPPC of TC. (I) has cytostatic activity and can be
XX used in vaccine production and as a tumour-specific immunogenic response
XX inducer. (I) is useful for treating 71 types of cancers or tumours in a
XX subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
XX oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
XX ectodermal tumour (PNET). (I) is useful as cancer immunogen including
XX vaccines. (I) is useful for diagnostic and palliative use, for detecting
XX or imaging cancer cells, and to monitor the course of amelioration of
XX malignancy in an individual. AAM43707 to AAM47109 represent peptides
XX which are used in the exemplification of the present invention
XX
XX Sequence 7 AA;
XX
XX Query Match 100.0%; Score 22; DB 4; Length 7;
XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

XX
XX 25-OCT-2001 (first entry)
XX
XX H11 binding site consensus conforming peptide (CCP) #2169.
XX
XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
XX immunogenically cross-reactive; cancer; immunogenic cancer cell;
XX cytostatic; vaccine; tumour-specific immunogenic response inducer;
XX astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
XX ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX CA2290722-A1.
XX
XX 08-JUN-2001.
XX
XX 08-DEC-1999; 99CA-02290722.
XX
XX 08-DEC-1999; 99CA-02290722.
XX
XX 08-DEC-1999; 99CA-02290722.
XX
XX (NOVO-) NOVOPHARM BIOTECH INC.
XX
XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
XX Entwistle JM, Macdonald GC;
XX
XX WPI; 2001-425937/46.
XX
XX Composition useful for treating and diagnosing cancer; comprises stress
XX protein-peptide complexes associated with tumor, and isolated antigen-
XX binding fragments of an antibody that binds specifically to the complex.
XX
XX Example 4; Page 108; 154pp; English.
XX
XX The present invention describes a composition (I) comprising stress
XX protein-peptide complexes (SPPC) associated with tumours that is
XX specifically immunogenically cross-reactive with cell surface-associated
XX SPPCs specific to target cancer (TC). Also described is an isolated
XX antigen-binding fragment of an antibody that binds specifically to SPPCs
XX or a population of different SPPCs consisting of immunogenic cancer cell
XX surface-associated SPPC of TC. (I) has cytostatic activity and can be
XX used in vaccine production and as a tumour-specific immunogenic response
XX inducer. (I) is useful for treating 71 types of cancers or tumours in a
XX subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
XX oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
XX ectodermal tumour (PNET). (I) is useful as cancer immunogen including
XX vaccines. (I) is useful for diagnostic and palliative use, for detecting
XX or imaging cancer cells, and to monitor the course of amelioration of
XX malignancy in an individual. AAM43707 to AAM47109 represent peptides
XX which are used in the exemplification of the present invention
XX
XX Sequence 7 AA;
XX
XX Query Match 100.0%; Score 22; DB 4; Length 7;
XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 KKFg 4
XX 2 KKFg 5
XX
XX RESULT 8
XX AAM4584
XX ID AAM44584 standard; peptide; 7 AA.
XX
XX AAM44584;
XX
XX 25-OCT-2001 (first entry)
XX
XX H11 binding site consensus conforming peptide (CCP) #855.
XX

```

KM Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
 KM immunogenically cross-reactive; cancer; immunogenic cancer cell;
 KM cytostatic; vaccine; tumour-specific immunogenic response inducer;
 KM astrocytoma; fibrosarcoma; myosarcoma; liposarcoma; oligodendroglioma;
 KM ependymoma; medulloblastoma; primitive neural ectodermal tumour.
 XX
 OS Homo sapiens.
 OS Synthetic.
 PN CA2290722-A1.
 XX
 PD 08-JUN-2001.
 XX
 PF 08-DEC-1999; 99CA-02290722.
 XX
 PR 08-DEC-1999; 99CA-02290722.
 XX
 PA (NOVO-) NOVOPHARM BIOTECH INC.
 PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
 PI Entwistle JM, Macdonald GC;
 PT WPI; 2001-425937/46.
 XX
 PS Composition useful for treating and diagnosing cancer, comprises stress
 PT protein-peptide complexes associated with tumor, and isolated antigen-
 PT binding fragments of an antibody that binds specifically to the complex.
 XX
 PS Example 4; Page 104; 154pp; English.
 XX
 CC The present invention describes a composition (I) comprising stress
 CC protein-peptide complexes (SPPC) associated with tumours that is
 CC specifically immunogenically cross-reactive with cell surface-associated
 CC SPPCs specific to target cancer (TC). Also described is an isolated
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs
 CC or a population of different SPPCs consisting of immunogenic cancer cell
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
 CC used in vaccine production and as a tumour-specific immunogenic response
 CC inducer. (II) is useful for treating 71 types of cancers or tumours in a
 CC subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
 CC or imaging cancer cells, and to monitor the course of amelioration of
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
 CC which are used in the exemplification of the present invention
 CC
 SQ Sequence 7 AA;
 QY
 Db 1 KKFG 4
 2 KKFG 5
 Query Match 100.0%; Score 22; DB 4; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX
 OS Homo sapiens.
 OS Synthetic.
 PN CA2290722-A1.
 XX
 PD 08-JUN-2001.
 XX
 PF 08-DEC-1999; 99CA-02290722.
 XX
 PR 08-DEC-1999; 99CA-02290722.
 XX
 PA (NOVO-) NOVOPHARM BIOTECH INC.
 PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
 PI Entwistle JM, Macdonald GC;
 PT WPI; 2001-425937/46.
 XX
 PS Composition useful for treating and diagnosing cancer, comprises stress
 PT protein-peptide complexes associated with tumor, and isolated antigen-
 PT binding fragments of an antibody that binds specifically to the complex.
 XX
 PS Example 4; Page 108; 154pp; English.
 XX
 CC The present invention describes a composition (I) comprising stress
 CC protein-peptide complexes (SPPC) associated with tumours that is
 CC specifically immunogenically cross-reactive with cell surface-associated
 CC SPPCs specific to target cancer (TC). Also described is an isolated
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs
 CC or a population of different SPPCs consisting of immunogenic cancer cell
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
 CC used in vaccine production and as a tumour-specific immunogenic response
 CC inducer. (II) is useful for treating 71 types of cancers or tumours in a
 CC subject, such as astrocytoma, fibrosarcoma, myosarcoma, liposarcoma,
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
 CC or imaging cancer cells, and to monitor the course of amelioration of
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
 CC which are used in the exemplification of the present invention
 CC
 SQ Sequence 7 AA;
 QY
 Db 1 KKFG 4
 2 KKFG 5
 Query Match 100.0%; Score 22; DB 4; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 08-JUN-2001.
 PD 08-DEC-1999; 99CA-02290722.
 XX 08-DEC-1999; 99CA-02290722.
 PF 08-DEC-1999; 99CA-02290722.
 XX (NOVO-) NOVOPHARM BIOTECH INC.
 XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
 PI Entwistle JM, MacDonald GC;
 XX WPI; 2001-425937/46.
 DR
 XX Composition useful for treating and diagnosing cancer, comprises stress
 PT protein-peptide complexes associated with tumor, and isolated antigen-
 PT binding fragments of an antibody that binds specifically to the complex.
 XX Example 4; Page 104; 154pp; English.
 PS
 XX The present invention describes a composition (I) comprising stress
 CC protein-peptide complexes (SPPC) associated with tumors that is
 CC specifically immunogenically cross-reactive with cell surface-associated
 CC SPSCs specific to target cancer (TC). Also described is an isolated
 CC antigen-binding fragment of an antibody that binds specifically to SPSCs
 CC or a population of different SPSCs consisting of immunogenic cancer cell
 CC surface-associated SPSC of TC. (I) has cytostatic activity and can be
 CC used in vaccine production and as a tumor-specific immunogenic response
 CC inducer. (I) is useful for treating 71 types of cancers or tumors in a
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
 CC ectodermal tumor (PNET). (I) is useful as cancer immunogen including
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
 CC or imaging cancer cells, and to monitor the course of amelioration of
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
 CC which are used in the exemplification of the present invention
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 22; DB 4; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KKFg 4
 Db 2 KKFg 5
 XX
 RESULT 11
 AAB35204
 ID AAB35204 standard; peptide; 8 AA.
 XX
 AC AAB35204;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Retroviral recombination assay peptide fragment #2.
 XX
 KM Retroviral recombination assay; gene therapy vector; viral vector; gag;
 KM pol; replication; HIV.
 XX
 OS Unidentified.
 XX
 PN WO200104360-A2.
 PD 18-JAN-2001.
 PF 06-JUL-2000; 2000WO-US018597.
 XX
 PR 09-JUL-1999; 99US-0143015P.
 PR 10-NOV-1999; 99US-0164626P.
 XX
 PA (UABR-) UAB RES FOUNDD.

PA (TRAN-) TRANZYME INC.
 XX Kappes JC, Wu X, Wakefield J;
 PI WPI; 2001-091927/10.
 DR N-PDB; AAF24357.
 XX
 PT Retroviral recombination assays, systems and cells, useful for evaluating
 PT the risk of producing a replication-competent retrovirus from a
 PT retroviral-based vector.
 XX
 PS Disclosure; Fig 2; 86pp; English.
 XX
 CC The present invention describes a method for detecting a retroviral
 CC genetic recombinant having gag and pol functions, involving a cell
 CC suspected of having a recombinant and propagating the recombinant in the
 CC presence of any necessary helper functions. This is useful as it enables
 CC the detection of vectors which are able to replicate in the host cell,
 CC and allows the production of vectors suitable for gene therapy. The
 CC present sequence is a peptide fragment used to demonstrate the method of
 CC the invention
 XX
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 22; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KKFg 4
 Db 5 KKFg 8
 XX
 RESULT 12
 AAG84320
 ID AAG84320 standard; peptide; 10 AA.
 XX
 AC AAG84320;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Arabidopsis thaliana peptide ligand #960.
 XX
 KM Plant; peptide pesticide; peptide herbicide; agricultural research.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200142279-A2.
 PD 14-JUN-2001.
 PF 13-DEC-2000; 2000WO-GB004781.
 XX
 PR 13-DEC-1999; 99GB-00029469.
 XX
 PA (PROT-) PROTEOM LTD.
 XX
 PI Roberts GM, Heal JR;
 XX WPI; 2001-381629/40.
 DR
 XX A set of peptide ligands for agricultural research and development of
 PT therapeutic agents comprise specific complementary peptides to proteins
 PT encoded by genes of plant genomes.
 XX
 PS Example 4; Page 165; 201pp; English.
 XX
 CC The present invention relates to a set of peptide ligands consisting of
 CC specific complementary peptides to proteins encoded by genes of plant
 CC genomes. The present sequence is one such peptide from Arabidopsis
 CC thaliana. The peptides of the present invention are useful in an assay to
 CC identify a peptide, especially a peptide pesticide or herbicide. The
 CC peptides are also useful for tools for agricultural research and

CC development
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 22; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 1 KKF 4

RESULT 13
AAG6341
ID AAG6341 standard; peptide; 10 AA.

AC AAG6341;
XX 18-SEP-2001 (first entry)

DE Human complementary peptide, SEQ ID NO: 2535.

KW Human; complementary peptide; ligand; drug discovery; drug design.

OS Homo sapiens.

PN WO200142277-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB004776.

PR 13-DEC-1999; 99GB-00029464.

PA (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
XX candidates or pro-drugs.

PS Example 4; Page 404; 646pp; English.

XX The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification

SQ Sequence 10 AA;

Query Match 100.0%; Score 22; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 3 KKF 6

RESULT 14
AAG6245
ID AAG6245 standard; peptide; 10 AA.

AC AAG6245;
XX 18-SEP-2001 (first entry)

XX Human complementary peptide, SEQ ID NO: 2439.

KW Human; complementary peptide; ligand; drug discovery; drug design.

OS Homo sapiens.

PN WO200142277-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB004776.

PR 13-DEC-1999; 99GB-00029464.

PA (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
XX candidates or pro-drugs.

PS Example 4; Page 391; 646pp; English.

XX The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification

SQ Sequence 10 AA;

Query Match 100.0%; Score 22; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 5 KKF 8

RESULT 15
AAG6183
ID AAG6183 standard; peptide; 10 AA.

AC AAG6183;
XX 11-SEP-2001 (first entry)

DE Saccharomyces cerevisiae peptide, SEQ ID NO: 1132.

KW Saccharomyces cerevisiae; complementary peptide; peptide identification;

KW drug discovery; drug design.

OS Saccharomyces cerevisiae.

PN WO200142276-A1.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB004773.

PR 13-DEC-1999; 99GB-00029471.

PA (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI; 2001-367863/38.

XX Identifying complementary peptides by analysis of protein and nucleotide
PT sequence databases, useful in drug design.

XX Example 3; Page 186; 488pp; English.

XX The invention relates to the identification of complementary peptides by
CC analysis of protein and nucleotide sequence databases from higher
CC eukaryotic genomes, excluding human and plants. The specific
CC complementary peptides interact with their relevant target proteins
CC encoded in the eukaryote genome. The peptides may be used as reagents and
CC drugs for drug discovery and as lead ligands for drug design and
CC development. The present sequence is a complementary peptide from
CC Saccharomyces cerevisiae

XX Sequence 10 AA;

Query Match 100.0%; Score 22; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
|||
Db 2 KKFG 5

Search completed: March 18, 2004, 11:16:22
Job time : 4.65427 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:13:05 ; Search time 0.463895 Seconds
(without alignments)
445.152 Million cell updates/sec

Title: US-10-076-421-3

Perfect score: 22

Sequence: 1 KKFQ 4

Scoring table: BLOSUM62
Gapop 10.0, Gapept 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents: AA: *
1: /cgn2_6/ptodata/2/1aa/5A COMB .pep: *
2: /cgn2_6/ptodata/2/1aa/5B COMB .pep: *
3: /cgn2_6/ptodata/2/1aa/6A COMB .pep: *
4: /cgn2_6/ptodata/2/1aa/6B COMB .pep: *
5: /cgn2_6/ptodata/2/1aa/PTUS COMB .pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1 .pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	12	1	US-08-260-582-15
2	22	100.0	12	5	PCT-US95-05471-15
3	22	100.0	13	1	US-08-465-325-135
4	22	100.0	13	4	US-08-115-737-135
5	22	100.0	15	2	US-08-338-882-5
6	22	100.0	16	4	US-08-914-999-20
7	22	100.0	17	1	US-08-132-767-11
8	22	100.0	18	1	US-08-233-203-10
9	22	100.0	18	2	US-08-338-882-18
10	22	100.0	18	2	US-08-338-882-20
11	22	100.0	19	1	US-08-132-767-8
12	22	100.0	19	1	US-08-132-767-10
13	22	100.0	20	1	US-07-711-183D-6
14	22	100.0	20	1	US-07-908-455A-74
15	22	100.0	20	1	US-08-132-767-7
16	22	100.0	20	1	US-08-434-120-11
17	22	100.0	20	1	US-08-465-325-11
18	22	100.0	20	3	US-08-782-997A-30
19	22	100.0	20	4	US-09-115-737-11
20	22	100.0	20	6	5254535-11
21	22	100.0	21	1	US-07-711-183D-5
22	22	100.0	21	1	US-07-965-663A-1
23	22	100.0	21	1	US-07-965-663A-2
24	22	100.0	21	1	US-07-965-663A-3
25	22	100.0	21	1	US-07-965-663A-22
26	22	100.0	21	1	US-07-908-455A-73
27	22	100.0	21	1	US-08-132-767-6

28	22	100.0	21	1	US-08-132-767-9	Sequence 9, Appl1
29	22	100.0	21	1	US-08-434-120-10	Sequence 10, Appl1
30	22	100.0	21	1	US-08-465-325-10	Sequence 10, Appl1
31	22	100.0	21	2	US-08-448-600-9	Sequence 9, Appl1
32	22	100.0	21	3	US-08-937-610-17	Sequence 17, Appl1
33	22	100.0	21	4	US-09-115-737-10	Sequence 10, Appl1
34	22	100.0	21	6	5254535-10	Patent No. 5254535
35	22	100.0	22	1	US-07-711-183D-3	Sequence 3, Appl1
36	22	100.0	22	1	US-07-965-663A-6	Sequence 4, Appl1
37	22	100.0	22	1	US-07-965-663A-19	Sequence 6, Appl1
38	22	100.0	22	1	US-07-965-663A-19	Sequence 7, Appl1
39	22	100.0	22	1	US-07-965-663A-20	Sequence 19, Appl1
40	22	100.0	22	1	US-07-965-663A-20	Sequence 20, Appl1
41	22	100.0	22	1	US-07-965-663A-21	Sequence 21, Appl1
42	22	100.0	22	1	US-07-908-455A-71	Sequence 71, Appl1
43	22	100.0	22	1	US-07-908-455A-72	Sequence 72, Appl1
44	22	100.0	22	1	US-08-132-767-5	Sequence 5, Appl1
45	22	100.0	22	1	US-08-132-767-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-08-260-582-15
; Sequence 15, Application US/08260582
; Patent No. 5635182
; GENERAL INFORMATION:
; APPLICANT: McCoy, John M.
; APPLICANT: Lu, Zhijian
; TITLE OF INVENTION: METHOD OF DETECTING LIGAND
; TITLE OF INVENTION: INTERACTIONS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/260,582
; FILING DATE: 16-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meinerdt, M. C.
; REGISTRATION NUMBER: 31,544
; REFERENCE/DOCKET NUMBER: GI 5236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-260-582-15

Query Match 100.0%; Score 22; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 4; Conservative 0; Mismatches 0; Gaps 0;
1 KKFQ 4
||||

Db 6 KKFG 9

RESULT 2
PCT-US95-05471-15
Sequence 15, Application PC/TUS9505471
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHOD OF DETECTING LIGAND INTERACTIONS
NUMBER OF SEQUENCES: 76
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION FOR SEQ ID NO: 15:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEITICAL: NO
ANTI-SENSE: NO
PCT-US95-05471-15

Query Match 100.0%; Score 22; DB 5; Length 12;
Best local Similarity 100.0%; Pred. No. 53;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFG 4
Db 6 KKFG 9

RESULT 3
US-08-465-325-135
Sequence 135, Application US/08465325
Patent No. 5686563
GENERAL INFORMATION:
APPLICANT: Magalini Pharmaceuticals Inc.
APPLICANT: 5110 Campus Drive
APPLICANT: Plymouth Meeting, PA 19462
TITLE OF INVENTION: Biologically Active Peptides Having
TITLE OF INVENTION: N-Terminal Substitutions
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunner
STREET: 1300 I. Street, N.W. Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,325
FILING DATE: 05-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/184,462
FILING DATE: 18-JAN-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/891,201
FILING DATE: 01-JUN-92
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B

REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0021-03000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 135:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-465-325-135

Query Match 100.0%; Score 22; DB 1; Length 13;
Best local Similarity 100.0%; Pred. No. 57;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFG 4
Db 1 KKFG 4

RESULT 4
US-09-115-737-135
Sequence 135, Application US/09115737
Patent No. 6348445
GENERAL INFORMATION:
APPLICANT: U. Prasad Kari
Taffy J. Williams
Michael McLane
TITLE OF INVENTION: Biologically Active Peptides With Reduced
Toxicity in Animals and a Method for Preparing Same
NUMBER OF SEQUENCES: 156
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunner, L.L.P.
STREET: 1300 I Street, N.W. Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/115,737
FILING DATE: 15-Jul-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,330
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: 08/184,462
FILING DATE: 18-JAN-94
APPLICATION NUMBER: 07/891,201
FILING DATE: 01-JUN-92
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0021-06000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 135:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 135:

US-09-115-737-135

Query Match 100.0%; Score 22; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
11111
DB 1 KKEG 4

RESULT 5

US-08-338-882-5

Sequence 5, Application US/0833882

Patent No. 5912231

GENERAL INFORMATION:

APPLICANT: Houghten, Richard A.

APPLICANT: Cuervo, Julio H.

TITLE OF INVENTION: Substitution Analogues of Maganin

TITLE OF INVENTION: Peptides

NUMBER OF SEQUENCES: 58

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dressler, Goldsmith, Shore, Suter &

ADDRESSEE: Milamow, Ltd.

STREET: 180 No. 5912231th Stetson, Suite 4700

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/338,882

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/615,125

FILING DATE: 15-NOV-1990

ATTORNEY/AGENT INFORMATION:

NAME: Gansson, Edward P.

REGISTRATION NUMBER: 29,381

REFERENCE/DOCKET NUMBER: SCRF 186.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 616-5400

TELEFAX: (312) 616-5460

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified-site

LOCATION: 15

OTHER INFORMATION: /note= "Xaa=alanine amide"

US-08-338-882-5

Query Match 100.0%; Score 22; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
11111
DB 10 KKEG 13

RESULT 6

US-08-914-999-20

Sequence 20, Application US/08914999

Patent No. 6346406

GENERAL INFORMATION:

APPLICANT: Ryazanov, Alexey G.

APPLICANT: Halt, William N.

APPLICANT: Pavur, Karen S.

TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)

TITLE OF INVENTION: AND METHODS OF USE THEREFOR

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.

STREET: 411 Hackensack Ave, Continental Plaza, 4th

STREET: Floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/914,999

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 601-1-078

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-345-1684

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-914-999-20

Query Match 100.0%; Score 22; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
11111
DB 2 KKEG 5

RESULT 7

US-08-132-767-11

Sequence 11, Application US/08132767

Patent No. 5518912

GENERAL INFORMATION:

APPLICANT: Michael A. Zaslloff, Nicole Resnick

TITLE OF INVENTION: NOVEL ENDOPEPTIDASE

NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5518912ris

STREET: One Liberty Place - 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: WORDPERFECT 5.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/132,767

FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/002,109
FILING DATE:
APPLICATION NUMBER: US/07/685,723
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Patricia A. Schreck
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: CH-0006
TELEPHONE: (215) 568-3439
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 17
TYPE: Amino acids
TOPOLOGY: Unknown
US-08-132-767-11

Query Match 100.0%; Score 22; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
Db 10 KKFg 13

RESULT 8
US-08-233-203-10

Sequence 10, Application US/08233203
Patent No. 5409898
GENERAL INFORMATION:
APPLICANT: Darveau, Richard P.
APPLICANT: Blake, James J.
APPLICANT: Cosand, Wesley L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS CAUSED BY ORGANISM SENSITIVE TO BETA-LACTAM
TITLE OF INVENTION: ANTIBIOTICS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company, Patent
ADDRESS: Department
STREET: 3005 First Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,203
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/655,321
FILING DATE: 19-FEB-1991
APPLICATION NUMBER: US 07/484,020
FILING DATE: 23-FEB-1990
ATTORNEY/AGENT INFORMATION:
NAME: Poor, Brian W.
REGISTRATION NUMBER: 32,928
REFERENCE/DOCKET NUMBER: ON0063A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206/448-4775
TELEFAX: 206/448-4800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
US-08-233-203-10

Query Match 100.0%; Score 22; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
Db 15 KKFg 18

RESULT 9
US-08-338-882-18

Sequence 18, Application US/08338882
Patent No. 5912231
GENERAL INFORMATION:
APPLICANT: Houghten, Richard A.
APPLICANT: Cuervo, Julio H.
TITLE OF INVENTION: Substitution Analogues of Magainin
TITLE OF INVENTION: Peptides
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
ADDRESSEE: Milnamow, Ltd.
STREET: 180 No. 5912231th Stetson, Suite 4700
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,882
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/615,125
FILING DATE: 15-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamsen, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: SCRF 186.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5460
TELEFAX: (312) 616-5460
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-338-882-18

Query Match 100.0%; Score 22; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKFg 4
Db 6 KKFg 9

RESULT 10
US-08-338-882-20
Sequence 20, Application US/08338882
Patent No. 5912231

GENERAL INFORMATION:
APPLICANT: Houghten, Richard A.
APPLICANT: Cuervo, Julio H.
TITLE OF INVENTION: Substitution Analogues of Maganin
TITLE OF INVENTION: Peptides
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Shore, Suter &
ADDRESSEE: Milamow, Ltd.
STREET: 180 No. 5912231th Stetson, Suite 4700
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,882
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/615,125
FILING DATE: 15-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gansson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: SCRF 186.1
TELEPHONE: (312) 616-5460
TELEFAX: (312) 616-5460
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-338-882-20

Query Match 100.0%; Score 22; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 6 KKF 9

RESULT 11
US-08-132-767-8
Sequence 8, Application US/08132767
Patent No. 5518912
GENERAL INFORMATION:
APPLICANT: Michael A. Zasloff; Nicole Resnick
TITLE OF INVENTION: NOVEL ENDOPEPTIDASE
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5518912ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/132,767
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/002,109
FILING DATE:
APPLICATION NUMBER: US/07/685,723
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Patricia A. Schreck
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: CH-0006
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 19
TYPE: Amino acids
TOPOLOGY: Unknown
US-08-132-767-8

Query Match 100.0%; Score 22; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
Db 6 KKF 9

RESULT 12
US-08-132-767-10
Sequence 10, Application US/08132767
Patent No. 5518912
GENERAL INFORMATION:
APPLICANT: Michael A. Zasloff; Nicole Resnick
TITLE OF INVENTION: NOVEL ENDOPEPTIDASE
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5518912ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/132,767
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/002,109
FILING DATE:
APPLICATION NUMBER: US/07/685,723
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Patricia A. Schreck
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: CH-0006
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19
TYPE: Amino acids
TOPOLOGY: Unknown
US-08-132-767-10

Query Match 100.0%; Score 22; DB 1; Length 19;

Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
Db 10 KKFG 13

RESULT 13
US-07-711-183D-6
; Sequence 6, Application US/07711183D
; Patent No. 5254537
; GENERAL INFORMATION:
; APPLICANT: Zaslloff, Michael
; TITLE OF INVENTION: Composition and Treatment with
; TITLE OF INVENTION: Peptide Combinations
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/711,183D
; FILING DATE: 19910510
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07 346894
; FILING DATE: 03-MAY-1989
; APPLICATION NUMBER: US 07 302985
; FILING DATE: 30-JAN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 421250-29
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 6
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: magainin peptide.
; PUBLICATION INFORMATION:
; AUTHORS: Zaslloff, Michael
; JOURNAL: Proceedings of the National Academy
; OF SCIENCES
; VOLUME: 84
; PAGES: 5449-5453
; DATE: AUG - 1987
; DOCUMENT NUMBER: US 4810777
; FILING DATE: 04-MAR-1987
; PUBLICATION DATE: 07-MAR-1989
; US-07-711-183D-6

Query Match 100.0%; Score 22; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
Db 7 KKFG 10

RESULT 14
US-07-908-455A-74
; Sequence 74, Application US/07908455A
; Patent No. 5459237
; GENERAL INFORMATION:
; APPLICANT: Berkowitz, Barry A.
; APPLICANT: Kari, U. Prasad
; APPLICANT: Maloy, W. Lee
; TITLE OF INVENTION: No. 5459237el Peptide Compositions and
; TITLE OF INVENTION: Uses therefor
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,455A
; FILING DATE: 19920702
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07686115
; FILING DATE: 15-APR-1991
; APPLICATION NUMBER: US 07476629
; FILING DATE: 08-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 421250-122
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: magainin peptide.
; OTHER INFORMATION: amide- or carboxy- terminated
; PUBLICATION INFORMATION:
; AUTHORS: Zaslloff, Michael
; JOURNAL: Proc. Nat. Acad. Sci.
; VOLUME: 84
; PAGES: 5449-5453
; DATE: AUG - 1987
; DOCUMENT NUMBER: US 4810777
; FILING DATE: 04-MAR-1987
; PUBLICATION DATE: 07-MAR-1989
; US-07-908-455A-74

Query Match 100.0%; Score 22; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFG 4
Db 7 KKFG 10

RESULT 15

US-08-132-767-7
; Sequence 7, Application US/08132767
; Patent No. 5518912
; GENERAL INFORMATION:
; APPLICANT: Michael A. Zaeloff; Nicole Resnick
; TITLE OF INVENTION: NOVEL ENDOPEPTIDASE
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5518912r1s
; STREET: One Liberty Place - 46th Floor
; City: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/132,767
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/002,109
; FILING DATE:
; APPLICATION NUMBER: US/07/685,723
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Patricia A. Schreck
; REGISTRATION NUMBER: 33,777
; REFERENCE/DOCKET NUMBER: CH-0006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: Amino acids
; TOPOLOGY: Unknown
; US-08-132-767-7

Query Match 100.0%; Score 22; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
|||
Db 7 KKEG 10

Search completed: March 18, 2004, 11:21:22
Job time: 1.46389 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:19:25 ; Search time 1.19037 Seconds
(without alignments)
870.166 Million cell updates/sec

Title: US-10-076-421-3

Perfect score: 22

Sequence: 1 KFRG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	4	13 US-10-076-421-3	Sequence 3, Appli
2	22	100.0	10	10 US-09-572-404B-2439	Sequence 2439, Ap
3	22	100.0	10	10 US-09-572-404B-2535	Sequence 2535, Ap
4	22	100.0	10	10 US-09-572-270A-960	Sequence 960, App
5	22	100.0	10	10 US-09-573-823C-328	Sequence 328, App
6	22	100.0	16	9 US-09-994-485-20	Sequence 20, Appli
7	22	100.0	22	9 US-09-807-120-3	Sequence 3, Appli
8	22	100.0	22	10 US-09-820-053A-24	Sequence 24, Appli
9	22	100.0	22	10 US-09-904-753-3	Sequence 3, Appli
10	22	100.0	22	10 US-09-904-753-4	Sequence 4, Appli
11	22	100.0	22	14 US-10-109-171-24	Sequence 24, Appli
12	22	100.0	23	9 US-09-030-619-211	Sequence 211, App
13	22	100.0	23	10 US-09-820-053A-7	Sequence 7, Appli
14	22	100.0	23	10 US-09-904-753-2	Sequence 2, Appli
15	22	100.0	23	10 US-09-798-026B-14	Sequence 14, Appli

16	22	100.0	23	14 US-10-109-171-7	Sequence 7, Appli
17	22	100.0	23	14 US-10-252-773-13	Sequence 13, Appli
18	22	100.0	23	15 US-10-277-233-211	Sequence 211, App
19	22	100.0	38	12 US-10-424-599-165729	Sequence 165729,
20	22	100.0	41	12 US-10-424-599-270907	Sequence 270907,
21	22	100.0	43	14 US-10-097-111-475	Sequence 475, App
22	22	100.0	43	14 US-10-321-857-65	Sequence 65, Appli
23	22	100.0	43	14 US-10-318-675-65	Sequence 1, Appli
24	22	100.0	44	12 US-10-349-543-1	Sequence 144596,
25	22	100.0	46	12 US-10-424-599-144596	Sequence 144596,
26	22	100.0	46	12 US-10-424-599-173819	Sequence 223967,
27	22	100.0	46	12 US-10-424-599-223967	Sequence 236167,
28	22	100.0	46	12 US-10-424-599-236167	Sequence 2, Appli
29	22	100.0	47	9 US-09-880-503-2	Sequence 250270,
30	22	100.0	48	12 US-10-424-599-250270	Sequence 33477, A
31	22	100.0	49	9 US-09-864-761-33472	Sequence 208329,
32	22	100.0	49	12 US-10-424-599-208329	Sequence 5, Appli
33	22	100.0	49	14 US-10-349-543-5	Sequence 35403, A
34	22	100.0	50	9 US-09-864-761-35403	Sequence 272294,
35	22	100.0	50	12 US-10-424-599-272294	Sequence 1679, Ap
36	22	100.0	51	9 US-09-925-301-1679	Sequence 43, Appli
37	22	100.0	51	9 US-09-250-611-43	Sequence 249369,
38	22	100.0	51	12 US-10-424-599-249369	Sequence 283379,
39	22	100.0	51	12 US-10-424-599-283379	Sequence 170813,
40	22	100.0	52	12 US-10-424-599-170813	Sequence 996, App
41	22	100.0	53	11 US-09-833-245-996	Sequence 37535, A
42	22	100.0	54	9 US-09-864-761-37535	Sequence 31376, A
43	22	100.0	55	12 US-10-424-599-235683	Sequence 209467,
44	22	100.0	55	12 US-10-029-386-31376	
45	22	100.0	57	12 US-10-424-599-209467	

ALIGNMENTS

RESULT 1
US-10-076-421-3
; Sequence 3, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, MANABU
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-3

Query Match 100.0%; Score 22; DB 13; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KFRG 4
Db 1 KFRG 4
RESULT 2
US-09-572-404B-2439
; Sequence 2439, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd

```

; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2439
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 396-405 and may interact with Sequence
; OTHER INFORMATION: in this patent.
US-09-572-404B-2439

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
   |||
Db 5 KKF 8

RESULT 3
US-09-572-404B-2535
; Sequence 2535, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2535
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 398-407 and may interact with Sequence
; OTHER INFORMATION: in this patent.
US-09-572-404B-2535

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
   |||
Db 3 KKF 6

RESULT 4
US-09-572-270A-960
; Sequence 960, Application US/09572270A
; Publication No. US20030148368A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Inter- complementary peptide listing
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/572,270A
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 1144
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 960
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
; OTHER INFORMATION: Sequence located in RPS16. at 130-139 and may interact with
US-09-572-270A-960
```

```

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
   |||
Db 1 KKF 4

RESULT 5
US-09-573-822C-328
; Sequence 328, Application US/09573822C
; Publication No. US20030199011A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands generated from microbial genome seq
; FILE REFERENCE: Microbe patent
; CURRENT APPLICATION NUMBER: US/09/573,822C
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 804
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 328
; LENGTH: 10
; TYPE: PRT
; ORGANISM: mycoplasma genitalium
; FEATURE:
; OTHER INFORMATION: Sequence located in MG214 at 76-85 and may interact with Sequence
; OTHER INFORMATION: this patent.
US-09-573-822C-328

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
   |||
Db 1 KKF 4

RESULT 6
US-09-994-485-20
; Sequence 20, Application US/09994485
; Patent No. US20020142429A1
; GENERAL INFORMATION:
; APPLICANT: Ryazanov, Alexey G.
; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESS: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/994,485
; FILING DATE: 27-No. US20020142429A1-2001
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/914,999
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
```


REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-994-485-20

Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 16;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
2 KKF 5

Db 10 KKF 13

RESULT 7
US-09-807-720-3
Sequence 3, Application US/09807720
Patent No. US20020162135A1
GENERAL INFORMATION:
APPLICANT: DANIEL, HENRY
TITLE OF INVENTION: EXPRESSION OF AN ANTIMICROBIAL PEPTIDE VIA THE PLASTID
FILE REFERENCE: 1462-PCT-US-00
CURRENT APPLICATION NUMBER: US/09/807,720
CURRENT FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/185,662
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 22
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-807-720-3

Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 22;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
10 KKF 13

Db 10 KKF 13

RESULT 8
US-09-820-053A-24
Sequence 24, Application US/09820053A
Publication No. US20030083243A1
GENERAL INFORMATION:
APPLICANT: Owen, Donald R.
TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
FILE REFERENCE: HELIX27
CURRENT APPLICATION NUMBER: US/09/820,053A
CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 165
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 24
LENGTH: 22
TYPE: PRT
ORGANISM: ARTIFICIAL SEQUENCE

FEATURE:
OTHER INFORMATION: SYNTHETIC SEQUENCE
NAME/KEY: MOD RES
LOCATION: (22)
OTHER INFORMATION: AMIDATION
US-09-820-053A-24

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 22;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
10 KKF 13

Db 10 KKF 13

RESULT 9
US-09-904-753-3
Sequence 3, Application US/09904753
Publication No. US20030092612A1
GENERAL INFORMATION:
APPLICANT: Lynos, Robert T
TITLE OF INVENTION: Use of Antimicrobial Peptides as Preservatives in
TITLE OF INVENTION: Ophthalmic Preparations, Including Solutions,
Emulsions, and Suspensions
FILE REFERENCE: 2973 ver 2
CURRENT APPLICATION NUMBER: US/09/904,753
CURRENT FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: WO 96/25183
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 22
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (22)
OTHER INFORMATION: Xaa at position 22 is Lys-amide
US-09-904-753-3

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 22;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
10 KKF 13

Db 10 KKF 13

RESULT 10
US-09-904-753-4
Sequence 4, Application US/09904753
Publication No. US20030092612A1
GENERAL INFORMATION:
APPLICANT: Lynos, Robert T
TITLE OF INVENTION: Use of Antimicrobial Peptides as Preservatives in
TITLE OF INVENTION: Ophthalmic Preparations, Including Solutions,
Emulsions, and Suspensions
FILE REFERENCE: 2973 ver 2
CURRENT APPLICATION NUMBER: US/09/904,753
CURRENT FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: WO 96/25183
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 22
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: maginin analog
US-09-904-753-4

Query Match 100.0%; Score 22; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 11

US-10-109-171-24
; Sequence 24, Application US/10109171
; Publication No. US20030109452A1
; GENERAL INFORMATION:
; APPLICANT: Owen, Donald R.
; TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES AND METHODS FOR THEIR USE
; FILE REFERENCE: HELX028
; CURRENT APPLICATION NUMBER: US/10/109,171
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 22
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC SEQUENCE
; NAME/KEY: MOD_RES
; LOCATION: (22)
; OTHER INFORMATION: AMIDATION
US-10-109-171-24

Query Match 100.0%; Score 22; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 12

US-09-030-619-211
; Sequence 211, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erile, Douglas
; APPLICANT: Frazer, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 66081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 211
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-030-619-211

Query Match 100.0%; Score 22; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 13
US-09-820-053A-7
; Sequence 7, Application US/09820053A
; Publication No. US20030083243A1
; GENERAL INFORMATION:
; APPLICANT: Owen, Donald R.
; TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
; FILE REFERENCE: HELX027
; CURRENT APPLICATION NUMBER: US/09/820,053A
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 23
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC SEQUENCE
; NAME/KEY: MOD_RES
; LOCATION: (23)
; OTHER INFORMATION: AMIDATION
US-09-820-053A-7

Query Match 100.0%; Score 22; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 14
US-09-904-753-2
; Sequence 2, Application US/09904753
; Publication No. US20030092612A1
; GENERAL INFORMATION:
; APPLICANT: Lynos, Robert T
; TITLE OF INVENTION: Use of Antimicrobial Peptides as Preservatives in
; TITLE OF INVENTION: Ophthalmic Preparations, Including Solutions,
; TITLE OF INVENTION: Emulsions, and Suspensions
; FILE REFERENCE: 2973 ver 2
; CURRENT APPLICATION NUMBER: US/09/904,753
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: WO 96/25183
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-904-753-2

Query Match 100.0%; Score 22; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 10 KKF 13

RESULT 15
US-09-798-026B-14
; Sequence 14, Application US/09798026B
; Publication No. US20030148936A1
; GENERAL INFORMATION:

```

; APPLICANT: ALPHARMA AS
; TITLE OF INVENTION: Cytotoxic Peptides Modified by Bulky or Lipophilic Moieties
; FILE REFERENCE: 40745-2
; CURRENT APPLICATION NUMBER: US/09/798,026B
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: GB 9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 23
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(23)
; OTHER INFORMATION: Artificial peptide = native magainin peptide or mutated native
; OTHER INFORMATION: magainin peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)..(6)
; OTHER INFORMATION: position of possible substitution (L-->W)
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)..(8)
; OTHER INFORMATION: position of possible substitution (S-->W)
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (19)..(19)
; OTHER INFORMATION: position of possible substitution (E-->W)
; US-09-798-026B-14

```

```

Query Match          100.0%; Score 22; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 KKFG 4
        ||||
Db      10 KKFG 13

```

Search completed: March 18, 2004, 11:36:07
Job time: 1.44037 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:10:10 ; Search time 0.41179 Seconds
(without alignments)
935.309 Million cell updates/sec

Title: US-10-076-421-3

Perfect score: 22

Sequence: 1 KKF 4

Scoring table: BLOSUM62
Gap 10.0, Gape 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_78: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	37	1 HSWT93	histone H2A.3 - wh
2	22	100.0	43	2 A39313	chaperonin, 58K -
3	22	100.0	50	2 E95014	hypothetical prote
4	22	100.0	50	2 G97887	orf151 (imported) -
5	22	100.0	58	2 H83721	hypothetical prote
6	22	100.0	61	1 R3YM14	ribosomal protein
7	22	100.0	61	2 D98277	hypothetical prote
8	22	100.0	62	2 A65045	hypothetical prote
9	22	100.0	67	2 AE3277	LEU ribosomal prot
10	22	100.0	83	2 A46930	teg292 protein - m
11	22	100.0	84	2 G81425	50S ribosomal prot
12	22	100.0	84	2 E90005	ribosomal protein
13	22	100.0	86	2 F97061	hypothetical prote
14	22	100.0	87	2 F69542	conserved hypotet
15	22	100.0	88	2 A64557	ribosomal protein
16	22	100.0	88	2 E71951	ribosomal protein
17	22	100.0	88	2 T44265	ribosomal protein
18	22	100.0	88	2 T43610	probable IS1617 tr
19	22	100.0	89	2 AG2918	50S ribosomal prot
20	22	100.0	89	2 B97693	ribosomal protein
21	22	100.0	89	2 E87288	ribosomal protein
22	22	100.0	89	2 T03669	reverse transcript
23	22	100.0	90	2 T03671	reverse transcript
24	22	100.0	91	2 E75560	ribosomal protein
25	22	100.0	91	2 T23795	hypothetical prote
26	22	100.0	92	2 T03615	retrovirus-related
27	22	100.0	92	2 T03666	reverse transcript
28	22	100.0	94	2 S77047	transposase sst17
29	22	100.0	99	2 T32862	hypothetical prote

30	22	100.0	100	2 T11953	ribosomal protein
31	22	100.0	100	2 E69460	conserved hypotet
32	22	100.0	107	2 T01742	hypothetical 12.6K
33	22	100.0	108	2 AG3005	conserved hypotet
34	22	100.0	111	2 T13343	hypothetical prote
35	22	100.0	115	2 T40621	hypothetical prote
36	22	100.0	117	2 G64335	hypothetical prote
37	22	100.0	117	2 H84651	hypothetical prote
38	22	100.0	118	1 PSKFR3	phospholipase A2 (
39	22	100.0	118	1 PSKFR2	phospholipase A2 (
40	22	100.0	118	1 PSKFR1	phospholipase A2 (
41	22	100.0	119	2 S74925	transposase s11065
42	22	100.0	119	2 S75488	transposase s1x211
43	22	100.0	119	2 S74836	transposase s1x085
44	22	100.0	119	2 S75827	transposase s1x128
45	22	100.0	119	2 S75590	transposase s11125

ALIGNMENTS

RESULT 1

HSWT93 histone H2A.3 - wheat (fragment)
C/Species: Triticum aestivum (common wheat)
C/Date: 31-May-1980 #sequence_revision 31-May-1980 #text_change 16-Feb-1997
C/Accession: A02604
R/Rodriguez, J.A.; Brandt, W.F.; von Holt, C.
Biochim. Biophys. Acta 578, 196-206, 1979
A/Title: Plant histone 2 from wheat germ, a family of histone H2A variants.
A/Accession: A90629; MUID:79209696; PMID:454655
A/Molecule type: protein
A/Residues: 1-37 <ROD>
A/Experimental source: germ
C/Superfamily: histone H2A
C/Keywords: acetylated amino end; chromosomal protein; DNA binding; nucleosome core
F/1/Modified site: acetylated amino end (Met) #status experimental

Query Match 100.0%; Score 22; DB 1; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.2e+02; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4

DB 12 KKF 15

RESULT 2

A39313 chaperonin, 58K - Thermus aquaticus (fragment)
C/Species: Thermus aquaticus
C/Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 12-Sep-1997
C/Accession: A39313
R/Raguchi, H.; Konishi, J.; Ishii, N.; Yoshida, M.
J. Biol. Chem. 266, 22411-22418, 1991
A/Title: A chaperonin from a thermophilic bacterium, Thermus thermophilus, that controls
A/Reference number: A39313; MUID:92042183; PMID:1662319
A/Accession: A39313
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-43 <TAG>
C/Superfamily: chaperonin groEL
C/Keywords: molecular chaperone

Query Match 100.0%; Score 22; DB 2; Length 43;

Best Local Similarity 100.0%; Pred. No. 1.4e+02; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4

DB 40 KKF 43

```
RESULT 3
E95014
hypothetical protein SP0124 [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C/Accession: E95014
R/Reteller: H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
neon, T.; Hickey, E.R.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Lofthus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A./Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: E95014
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-50 <KUR>
A/Cross-references: GB:AE005672; PIDN:AAK74310.1; PID:g14971592; GSPDB:GN00164; TIGR:SP4
A/Experimental source: strain TIGR4
A/Genetics:
A/Gene: SP0124

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 44 KKF 47

RESULT 4
G9787
orf51 [imported] - Streptococcus pneumoniae (strain R6)
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C/Accession: G9787
R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
y, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnhem, S.; M
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;
A./Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: G9787
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-50 <KUR>
A/Cross-references: GB:AE007317; PIDN:AAK98931.1; PID:g15457666; GSPDB:GN00174
A/Genetics:
A/Gene: orf51

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 44 KKF 47

RESULT 5
H83721
hypothetical protein BH0576 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: H83721
R/Takami, H.; Nakane, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A./Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: H83721

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-58 <STO>
A/Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BA04295.1; GSPDB:GN001
A/Experimental source: strain C-125
A/Genetics:
A/Gene: BH0576

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 58;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 30 KKF 33

RESULT 6
R3YM14
ribosomal protein S14 - Mycoplasma capricolum
C/Species: Mycoplasma capricolum
C/Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 07-Dec-1999
C/Accession: S02844
R/Ohkubo, S.; Muto, A.; Kawachi, Y.; Yamao, F.; Oosawa, S.
Mol. Gen. Genet. 210, 314-322, 1987
A./Title: The ribosomal protein gene cluster of Mycoplasma capricolum.
A/Reference number: S02830; MUID:88142549; PMID:3481422
A/Accession: S02844
A/Molecule type: DNA
A/Residues: 1-61 <OHK>
A/Cross-references: EMBL:X06414; NID:g44207; PIDN:CAA29717.1; PID:g44222
A/Genetics:
A/Gene: rps14
A/Gene code: SGC3
C/Superfamily: Escherichia coli ribosomal protein S14
C/Keywords: protein biosynthesis; ribosome

Query Match
Best Local Similarity 100.0%; Score 22; DB 1; Length 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 35 KKF 38

RESULT 7
D98277
hypothetical protein AGR_L_2351 [imported] - Agrobacterium tumefaciens (strain C58, Cerec
C/Species: Agrobacterium tumefaciens
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C/Accession: D98277
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A./Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tume
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: D98277
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-61 <KUR>
A/Cross-references: GB:AE007870; PIDN:AAK89742.1; PID:g15159660; GSPDB:GN00170
A/Genetics:
A/Gene: AGR_L_2351
A/Map position: linear chromosome

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 32 KKF 35
```

RESULT 8
A:Accession: A65045
hypothetical protein b2656 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: A65045
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: A65045
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-67 <BLAT>
A:Cross-references: GB:AE000350; GB:U00096; NID:92367147; PIDN:AACT5703.1; PID:gl789009;
A:Experimental source: strain K-12, substrain M01655

Query Match 100.0%; Score 22; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
Db 15 KKF 18

RESULT 9

AE3277
LST ribosomal protein L27 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 18-Nov-2002
C:Accession: AE3277
R:DelVecchio, V.G.; Kaputrl, V.; Redkar, R.U.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
A.; Mazur, M.; Goldsman, E.; Selkov, E.; Blizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3277
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL51384.1; PID:gl7982087; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME10202
A:Map position: 1
C:Superfamily: Escherichia coli ribosomal protein L27; eubacterial ribosomal protein L27

Query Match 100.0%; Score 22; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
Db 2 KKF 5

RESULT 10

S46930
teg292 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S46930
R:Lopez-Fernandez, L.A.; del Mazo, J.
submitted to the EMBL Data Library, March 1994
A:Reference number: S46929
A:Accession: S46930
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-83 <LOP>

A:Cross-references: EMBL:X60433; NID:9515352; PIDN:CAAS6628.1; PID:9393954

Query Match 100.0%; Score 22; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
Db 43 KKF 46

RESULT 11

G81425
50S ribosomal protein L27 Cj0095 [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Nov-2002
C:Accession: G81425
R:Perkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling-
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrett
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp-
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: G81425
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-84 <PAR>
A:Cross-references: GB:AL139074; GB:AL111168; NID:96967505; PIDN:CAB72579.1; PID:96967585
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: rplM, Cj0095
C:Superfamily: Escherichia coli ribosomal protein L27; eubacterial ribosomal protein L27

Query Match 100.0%; Score 22; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
Db 24 KKF 27

RESULT 12

E90005
ribosomal protein L31 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002
C:Accession: E90005
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc-
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; i
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: E90005
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-84 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701914; PIDN:BA43206.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: rpmE
C:Superfamily: Escherichia coli ribosomal protein L31

Query Match 100.0%; Score 22; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
Db 76 KKF 79

RESULT 13

F97061
hypothetical protein CAC1312 [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C/Accession: F97061
R/Molling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A/Reference number: A96900; MUID:2135325; PMID:21359325
A/Accession: F97061
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-86 <KIR>
A/Cross-references: GB:AE001437; PIDN:AAK79281.1; PID:GL5024241; GSPDB:GN00168
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CAC1312

Query Match 100.0%; Score 22; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

OY 1 KKFg 4
|||
Db 75 KKFg 78

RESULT 14
F69542
conserved hypothetical protein AF2342 - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C/Accession: F69542
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Ueberback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: F69542
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-87 <KLE>
A/Cross-references: GB:AE000943; GB:AE000782; NID:G2689266; PIDN:AA88914.1; PID:G264817
C/Superfamily: hypothetical protein MG0071

Query Match 100.0%; Score 22; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

OY 1 KKFg 4
|||
Db 23 KKFg 26

RESULT 15
A64557
ribosomal protein L27 - Helicobacter pylori (strain 26695)
C/Species: Helicobacter pylori
C/Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 13-Aug-1999
C/Accession: A64557
R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalaf, H.G.; Glodek, A.; McKen
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney, L.
Nature 388, 539-547, 1997
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A/Reference number: A64520; MUID:97394467; PMID:9252185
A/Accession: A64557
A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA
A/Residues: 1-88 <TOM>
A/Cross-references: GB:AE000548; GB:AE000511; NID:G2313391; PIDN:AA07366.1; PID:G2313391
C/Superfamily: Escherichia coli ribosomal protein L27; eubacterial ribosomal protein L27
F;2-84/Domain: eubacterial ribosomal protein L27 homology <L27>

Query Match 100.0%; Score 22; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

OY 1 KKFg 4
|||
Db 24 KKFg 27

Search completed: March 18, 2004, 11:20:17
Job time: 1.41138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:54:45 ; Search time 0.245077 Seconds

(without alignments)
849.859 Million cell updates/sec

Title: US-10-076-421-3

Perfect score: 22

Sequence: 1 KKEFG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.1*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	22	100.0	60	RS14_MYCCA	P10130 mycoplasma
2	22	100.0	61	RS14_CLOPE	Q8XHC6 clostridium
3	22	100.0	76	APM2_HUMAN	Q15847 homo sapien
4	22	100.0	84	R31B_PHOL	Q7NA98 photorhabd
5	22	100.0	84	R31B_STRAM	Q93688 streptococ
6	22	100.0	84	RL27_CAMJE	Q9P311 campylobact
7	22	100.0	85	R31B_STRAP	Q8CMA8 campylococ
8	22	100.0	88	RL27_HELPU	Q9ZM68 helicobacte
9	22	100.0	88	RL27_HELPU	P56050 helicobacte
10	22	100.0	88	RL27_AGRF5	Q93H66 agrobacteri
11	22	100.0	89	RL27_AGRF5	Q8VW58 bruceella ab
12	22	100.0	89	RL27_BRUM	Q8VW58 bruceella ab
13	22	100.0	89	RL27_BRUM	Q8VW58 bruceella me
14	22	100.0	89	RL27_BRUM	Q8VW58 bruceella su
15	22	100.0	89	RL27_CAVCR	Q9ABD3 caulobacter
16	22	100.0	89	RL27_RHIL	Q98E20 rhizobium 1
17	22	100.0	89	RL27_RHIL	Q921B7 rhizobium 1
18	22	100.0	91	RL27_CAEEL	Q21557 caenorhabdi
19	22	100.0	91	RL27_CAEEL	Q9RY65 deinoecoccus
20	22	100.0	96	Y12A_BP74	P37857 zamia fisch
21	22	100.0	100	RL27_CVACA	P39853 bacterioph
22	22	100.0	103	CHLB_ZAMFI	P37857 zamia fisch
23	22	100.0	105	DL4A_HUMAN	P96015 homo sapien
24	22	100.0	105	DL4A_MOUSE	Q9DCM4 mus musculu
25	22	100.0	111	RNPA_FUSNN	Q8RHA6 fusobacteri
26	22	100.0	115	PT19_STYPL	P28211 styela plic
27	22	100.0	117	Y286_MERTU	Q57734 methanococ
28	22	100.0	118	PA26_BUNFA	P00627 bungarus fa
29	22	100.0	118	PA26_BUNFA	P00627 bungarus fa
30	22	100.0	118	PA26_BUNFA	P00627 bungarus fa
31	22	100.0	119	RNPA_CHIMU	Q9P137 chlamydia m
32	22	100.0	120	RNPA_CHLTR	Q84789 chlamydia t
33	22	100.0	122	RL7_BUCAP	P41188 buchnera ap

34	22	100.0	122	1	RL7_BUCAP	Q89B19 buchnera ap
35	22	100.0	123	1	RL7_WIGBR	Q8D234 wigleswort
36	22	100.0	125	1	ACPS_NEIMA	Q9RM72 neisseria m
37	22	100.0	125	1	ACPS_NEIMA	Q9RM72 neisseria m
38	22	100.0	125	1	RL7_HELPU	P55834 helicobacte
39	22	100.0	130	1	RS8_MERTU	P54041 methanococ
40	22	100.0	131	1	RS9_MYCCA	Q7NB14 mycoplasma
41	22	100.0	132	1	RS9_MYCCE	P47656 mycoplasma
42	22	100.0	132	1	RS9_MYCCE	P75179 mycoplasma
43	22	100.0	133	1	RS9_UREPA	Q9PDC3 ureaplasma
44	22	100.0	134	1	RS9_MERTU	Q8TV55 methanopyru
45	22	100.0	135	1	RS16_PIG	Q29201 sus scrofa

ALIGNMENTS

RESULT 1
ID RS14_MYCCA STANDARD; PRT; 60 AA.
AC P10130;
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE 30S ribosomal protein S14.
GN RPSN.
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2095;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27343 / Kid;
RX MEDLINE=88142549; PubMed=3481422;
RA Okubo S., Muto A., Kawachi Y., Yamao F., Ogasawa S.;
RT "The ribosomal protein gene cluster of Mycoplasma capricolum.";
RL Mol. Genet. 210:314-322(1987).
CC -1- FUNCTION: Known to be required for the assembly of 30S particles
CC and may also be responsible for determining the conformation of
CC the 16S rRNA at the A site (By similarity).
CC -1- SIMILARITY: Belongs to the S14p family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X06414; CA29717.1; ALT_TERM.
DR PIR; S02844; R3YM14.
DR InterPro; IPR001209; Ribosomal_S14.
DR Pfam; PF00253; Ribosomal_S14; 1.
DR PROSITE; PS00527; RIBOSOMAL_S14; 1.
KW Ribosomal protein.
SQ SEQUENCE 60 AA; 6933 MW; 0409C5311B453CF0 CRC64;
Query Match Score 22; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKEFG 4
DB 35 KKEFG 36
RESULT 2
ID RS14_CLOPE STANDARD; PRT; 61 AA.
AC Q8XHC6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE 30S ribosomal protein S14.
GN RPSN OR CPB2392.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kihara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -1- FUNCTION: Known to be required for the assembly of 30S particles
CC and may also be responsible for determining the conformation of
CC the 16S rRNA at the A site (by similarity).
CC -1- SIMILARITY: Belongs to the S14P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AP003194; BAB82098.1; -
CC DR InterPro; IPR001209; Ribosomal_S14.
CC DR Pfam; PF00253; Ribosomal_S14; 1.
CC DR PROSITE; PS00527; RIBOSOMAL_S14; 1.
CC KM Ribosomal protein; Complete proteome.
CC SQ SEQUENCE 61 AA; 7280 MW; C3910B3227493E75 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 35 KKF 38

RESULT 3
APM2_HUMAN STANDARD; PRT; 76 AA.
ID APM2_HUMAN
AC Q15847;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adipose most abundant gene transcript 2.
GN APM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipose tissue;
RX MEDLINE=96224171; PubMed=8619847;
RA Maeda K., Okubo K., Shimomura I., Funahashi T., Matsuzawa Y.,
RA Matsubara K.;
RT "cDNA cloning and expression of a novel adipose specific collagen-like
RT factor, apm1 (Adipose Most abundant Gene transcript 1).";
RL Biochem. Biophys. Res. Commun. 221:286-289(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```

```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mallya S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Skalska U., Smalins D.E.,
RA Butlerfield Y.S.N., Krzywicki M.I., Skalska U.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THE ADIPOSE TISSUE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; DA5370; BAA08226.1; -
CC DR EMBL; BC004471; AAH04471.1; -
CC SQ SEQUENCE 76 AA; 7855 MW; C95B41F13C4BA7B1 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
DB 70 KKF 73

RESULT 4
R31B_PROL STANDARD; PRT; 84 AA.
ID R31B_PROL
AC Q7NA98;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 50S ribosomal protein L31 type B.
GN RPEM2 OR RPEM OR PU00036.
OS Photobacterium luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photobacterus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Bourreau-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freydisse G., Gaudriault S.,
RA Médigue C., Lanois A., Powell K., Siguer P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photobacterus
RT luminescens";
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -1- SIMILARITY: Belongs to the L31P family of ribosomal proteins.
CC Subfamily B.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; BX571859; CAB12331.1; -

DR Photolact; plu0036; -

DR HAMAP; MF_00502; -; 1.

DR PROSITE; PS01143; RIBOSOMAL_L31; FALSE_NEG.

KM Ribosomal protein; Complete proteome.

SO SEQUENCE 84 AA; 9593 MW; 39681B9E527105B2 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 84;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4

Db 75 KKF 78

RESULT 5

R31B STAM STANDARD; PRT; 84 AA.

ID R31B STAM

AC 0995DB;

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 505 ribosomal protein L31 type B.

GN RME2 OR RPME OR SAV2120 OR SAL1922 OR MW2044.

OS Staphylococcus aureus (strain M50 / ATCC 700699),

OS Staphylococcus aureus (strain N315), and

OS Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=158878, 158879, 196620;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=M50 / ATCC 700699, and N315;

RX MEDLINE=21311952; PubMed=11418146;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,

RA Kanamori M., Matsunaru H., Maruyama A., Murakami R., Hosoyama A.,

RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,

RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,

RA Kanetsu M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,

RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;

RT "Whole genome sequencing of methicillin-resistant Staphylococcus

RT aureus."

RL Lancet 357:1225-1240(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MW2;

RX MEDLINE=22040717; PubMed=12044378;

RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,

RA Nagai Y., Iwama N., Amano K., Naito T., Kuroda H., Cui L.,

RA Yamamoto K., Hiramatsu K.;

RT "Genome and virulence determinants of high virulence community-

RT acquired MRSA."

RL Lancet 359:1819-1827(2002).

RN [1]

RP SIMILARITY: Belongs to the L31P family of ribosomal proteins.

CC Subfamily B.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AP003364; BAB58282.1; -

DR EMBL; AP003136; BAB43206.1; -

DR EMBL; AP004629; BAB95909.1; -

DR PIR; E90005; E90005.

DR HAMAP; MF_00502; -; 1.

DR InterPro; IPR002150; Ribosomal_L31.

DR Pfam; PF01197; Ribosomal_L31; 1.

DR PRINTS; PRO1249; RIBOSOMAL_L31.

DR TIGRFAMs; TIGR00105; L31; 1.

DR PROSITE; PS01143; RIBOSOMAL_L31; 1.

KM Ribosomal protein; Complete proteome.

SO SEQUENCE 84 AA; 9723 MW; 28CA86AE4FDEE01 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 84;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4

Db 76 KKF 79

RESULT 6

RL27_CAMJ2 STANDARD; PRT; 84 AA.

ID RL27_CAMJ2

AC 0995J1;

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 505 ribosomal protein L27.

GN RPWA OR CU0095.

OS Campylobacter jejuni.

OS Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;

OC Campylobacteraceae; Campylobacter.

OX NCBI_TaxID=197;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 11168;

RX MEDLINE=20150912; PubMed=10688204;

RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,

RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,

RA Jorgensen K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,

RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,

RA Whitehead S., Barrall B.G.;

RT "The genome sequence of the food-borne pathogen Campylobacter jejuni

RT reveals hyper-variable sequences."

RL Nature 403:665-668(2000).

RN [1]

RP SIMILARITY: Belongs to the L27P family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AL139074; CAB72579.1; -

DR PIR; G81425; G81425.

DR HAMAP; MF_00539; -; 1.

DR InterPro; IPR001684; Ribosomal_L27.

DR Pfam; PR01016; Ribosomal_L27; 1.

DR PRINTS; PR00063; RIBOSOMAL_L27; 1.

DR PRODOM; PD003114; Ribosomal_L27; 1.

DR TIGRFAMs; TIGR00062; L27; 1.

DR PROSITE; PS00831; RIBOSOMAL_L27; 1.

KM Ribosomal protein; Complete proteome.

SO SEQUENCE 84 AA; 9285 MW; EEC60824EB1B0670 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 84;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4

Db 24 KKF 27

```

RESULT 7
R13B_STAMP STANDARD; PRT; 85 AA.
ID R13B_STAMP STANDARD; PRT; 85 AA.
AC O8CRM8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 50S ribosomal protein L31 type B.
GN RME2 OR SE1718.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qin Z.-Q., Mao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
CC -1- SIMILARITY: Belongs to the L31P family of ribosomal proteins.
CC Subfamily B.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE016749; AA005317.1; -.
DR HAMAP; MF_00502; -; 1.
DR InterPro; IPR002150; Ribosomal_L31.
DR Pfam; PF01197; Ribosomal_L31; 1.
DR PRINTS; PR01249; RIBOSOMAL_L31.
DR TIGRFAMs; TIGR00105; L31; 1.
DR PROSITE; PS01143; RIBOSOMAL_L31; 1.
DR Ribosomal protein; Complete proteome.
KW SEQUENCE 85 AA; 9779 MW; 65BBD6687DCA5AAD CRC64;
SQ
Query Match 100.0%; Score 22; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKFQ 4
Db 76 KKFQ 79

```

```

RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE001465; AAD05863.1; -.
DR PIR; E71951; E71951.
DR HAMAP; MF_00539; -; 1.
DR InterPro; IPR001684; Ribosomal_L27.
DR Pfam; PF01016; Ribosomal_L27; 1.
DR PRINTS; PR00063; RIBOSOMAL_L27.
DR ProDom; PD003114; Ribosomal_L27; 1.
DR TIGRFAMs; TIGR00062; L27; 1.
DR PROSITE; PS00831; RIBOSOMAL_L27; 1.
DR Ribosomal protein; Complete proteome.
KW SEQUENCE 88 AA; 9676 MW; 29B0AB6341D73B60 CRC64;
SQ
Query Match 100.0%; Score 22; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKFQ 4
Db 24 KKFQ 27

```

```

RESULT 8
RL27_HELPJ STANDARD; PRT; 88 AA.
ID RL27_HELPJ STANDARD; PRT; 88 AA.
AC O9ZMD8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPYA OR JHP0282.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deLonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uita-Nickelsen M., Mills D.W., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Yang Q., Taylor D.E., Voyis G.F.,
RA Trust T.J.;

```

```

RESULT 9
RL27_HELPJ STANDARD; PRT; 88 AA.
ID RL27_HELPJ STANDARD; PRT; 88 AA.
AC P56050;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPYA OR HP0297.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----

```

DR EMBL: A6000548; AAD07366.1; -
DR TIGR: A64557; A64557.
DR HAMAP: MF_00539; -; 1.
DR InterPro: IPR001684; Ribosomal_L27.
DR Pfam: PF01016; Ribosomal_L27; 1.
DR PRINTS: PR00063; RIBOSOMALL27.
DR PRODOM: PD003114; Ribosomal_L27; 1.
DR TIGRFAMs: TIGR00062; L27; 1.
DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
DR Riboosomal protein; Complete proteome.
SQ SEQUENCE 88 AA; 9778 MW; 29BDAE61A7951F0 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKF 4
Db 24 KKF 27

RESULT 10
RL27_SYN7 STANDARD; PRT; 88 AA.
AC Q923H6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPMA OR RPL27.
OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1140;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=9837326; PubMed=9727980.
RA Ishiura M., Kutsuna S., Aoki S., Iwasaki H., Andersson C.R.,
RA Tanabe A., Golden S.S., Johnson C.H., Kondo T.;
RT "Expression of a gene cluster kaiABC as a circadian feedback process
in cyanobacteria."
RL Science 281:1519-1523(1998).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB010691; BA37099.1; -
DR PIR: T44265; T44265.
DR HAMAP: MF_00539; -; 1.
DR InterPro: IPR001684; Ribosomal_L27.
DR Pfam: PF01016; Ribosomal_L27; 1.
DR PRINTS: PR00063; RIBOSOMALL27.
DR PRODOM: PD003114; Ribosomal_L27; 1.
DR TIGRFAMs: TIGR00062; L27; 1.
DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
DR Riboosomal protein.
SQ SEQUENCE 88 AA; 9226 MW; 0A448F819FFC42F0 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKF 4
Db 24 KKF 27

RESULT 11
RL27_AGR5 STANDARD; PRT; 89 AA.
AC Q8UBR6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPMA OR ATU2784 OR AGR_C_5052.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monke D.E., Kitaajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. Sr.,
RA Chapman P., Clendenning U., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58."
RL Science 294:2317-2323(2001).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Ourullo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houtmel K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,
RA Flanagan C., Crowell C., Gureon J., Lomo C., Sear C., Strub G.,
RA Cleo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB009226; AAL4365.1; -
DR PIR: AB008192; AAK88499.1; -
DR PIR: AG2918; AG2918.
DR HAMAP: MF_00539; -; 1.
DR InterPro: IPR001684; Ribosomal_L27.
DR Pfam: PF01016; Ribosomal_L27; 1.
DR PRINTS: PR00063; RIBOSOMALL27.
DR PRODOM: PD003114; Ribosomal_L27; 1.
DR TIGRFAMs: TIGR00062; L27; 1.
DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
DR Riboosomal protein; Complete proteome.
SQ SEQUENCE 89 AA; 9396 MW; 59D141BF2B47301A CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKF 4
Db 24 KKF 27

```

RESULT 12
RL27 BRUB STANDARD; PRT; 89 AA.
AC Q8VW58;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 50S ribosomal protein L27.
GN RPLM27.
OS Brucella abortus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21952364; PubMed=11955619;
RA Halling S.M., Zuercher R.L.;
RT "Evidence for lateral transfer to Brucellae: characterization of a
RT locus with a Tn-like element (Tn2020).";
RL Biochim. Biophys. Acta 1574:109-116(2002).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF119331; AAL32286.1; -
DR HAMAP; MF_00539; -1
DR InterPro; IPR001684; Ribosomal_L27.
DR Pfam; PF01016; Ribosomal_L27; 1.
DR PRINTS; PR00063; RIBOSOMALL27.
DR PRODOM; PD003114; Ribosomal_L27; 1.
DR TIGRFAMs; TIGR00062; L27; 1.
DR PROSITE; PS00831; RIBOSOMAL_L27; 1.
KW Ribosomal protein.
SQ SEQUENCE 89 AA; 9377 MW; 8667B61EC97BAD5 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFQ 4
DB 24 KKFQ 27

RESULT 13
RL27 BRUM STANDARD; PRT; 89 AA.
AC Q8YJ84;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L27.
GN RPLM27.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapatal V., Redkar R.J., Patra G., Muier C., Los T.,
RA Ivanova N., Andersen I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonaki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Teleson J.-J.,
RA Haselkorn R., Kyriades N., Overbeek R.;

```

```

RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014475; AAN30744.1; -
DR HAMAP; MF_00539; -1
DR InterPro; IPR001684; Ribosomal_L27.
DR Pfam; PF01016; Ribosomal_L27; 1.
DR PRINTS; PR00063; RIBOSOMALL27.
DR PRODOM; PD003114; Ribosomal_L27; 1.
DR TIGRFAMs; TIGR00062; L27; 1.
DR PROSITE; PS00831; RIBOSOMAL_L27; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 89 AA; 9407 MW; 3D22A8838C9CAD09 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFQ 4
DB 24 KKFQ 27

RESULT 14
RL27 BRUSU STANDARD; PRT; 89 AA.
AC Q8FYI8;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 50S ribosomal protein L27.
GN RPLM27.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Debroy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
RA Riedmiller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014475; AAN30744.1; -
DR TIGR; BR1849; -

```

DR HAMAP: MF 00539; -: 1.
DR InterPro: IPR001684; Ribosomal_L27.
DR Pfam: PF01016; Ribosomal_L27; 1.
DR PRINTS: PR00063; RIBOSOMAL_L27.
DR ProDom: PD003114; Ribosomal_L27; 1.
DR TIGRFAMs: TIGR00062; L27; 1.
DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 89 AA; 9377 MW; 8667B661EC87BAD5 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFg 4
DB 24 KKFg 27

RESULT 15
RL27_CAUCR STANDARD; PRT; 89 AA.
ID RL27_CAUCR
AC O9ABB3;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 50S ribosomal protein L27.
GN RPLM OR CC0318.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Ueberback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RA "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- SIMILARITY: Belongs to the L27P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: AE005705; AAK2305.1; -.
DR PIR: E87288; E87288.
DR TIGR: CC0318; -.
DR HAMAP: MF 00539; -: 1.
DR InterPro: IPR001684; Ribosomal_L27.
DR Pfam: PF01016; Ribosomal_L27; 1.
DR PRINTS: PR00063; RIBOSOMAL_L27.
DR ProDom: PD003114; Ribosomal_L27; 1.
DR TIGRFAMs: TIGR00062; L27; 1.
DR PROSITE: PS00831; RIBOSOMAL_L27; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 89 AA; 9371 MW; 22B85AECF5A28C91 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKFg 4

DB 24 KKFg 27

Search completed: March 18, 2004, 11:17:00
Job time: 1.24508 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:09:25 ; Search time 1.09409 Seconds
(without alignments)
1153.535 Million cell updates/sec

Title: US-10-076-421-3
Perfect score: 22
Sequence: 1 KKG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_Organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	25	10	Q94199 atropa bell
2	22	100.0	30	16	Q98MC2 rhizobium l
3	22	100.0	35	15	Q97875 human immun
4	22	100.0	50	16	Q97732 streptococc
5	22	100.0	50	16	Q8DRH5 streptococc
6	22	100.0	58	16	Q9KPA9 bacillus ha
7	22	100.0	59	5	Q96380 aedes albop
8	22	100.0	59	5	Q86PR6 culicx pipie
9	22	100.0	59	16	Q831G6 enterococcu
10	22	100.0	60	11	Q35588 mesocricetu
11	22	100.0	60	17	Q82Z07 pyrobaculum
12	22	100.0	61	16	Q8U4W9 agrobacteri
13	22	100.0	61	16	Q7VCU4 prochloroco
14	22	100.0	62	10	Q91OG1 arabidopsis
15	22	100.0	62	16	P7618 escherichia
16	22	100.0	63	16	Q8DL07 synechococc

17	22	100.0	64	10	Q9FRT4	Q9FRT4 brassica ol
18	22	100.0	65	16	Q8DB17	Q8DB17 vibrio vuln
19	22	100.0	65	16	Q81TK0	Q81TK0 bacillus an
20	22	100.0	66	9	Q85EB3	Q85EB3 mycobacteri
21	22	100.0	66	10	Q24089	Q24089 medicago tr
22	22	100.0	66	16	Q81LX1	Q81LX1 bacillus an
23	22	100.0	66	16	Q812T5	Q812T5 bacillus ce
24	22	100.0	67	12	Q91BL0	Q91BL0 spodoptera
25	22	100.0	67	13	Q93340	Q93340 struthio ca
26	22	100.0	68	16	Q87GC6	Q87GC6 vibrio para
27	22	100.0	69	16	Q8BEP6	Q8BEP6 streptococc
28	22	100.0	69	16	Q8B191	Q8B191 streptococc
29	22	100.0	69	16	Q87HL5	Q87HL5 vibrio para
30	22	100.0	69	16	Q81GFI	Q81GFI bacillus ce
31	22	100.0	70	10	Q9SXM3	Q9SXM3 lithospermu
32	22	100.0	70	10	Q9XHT1	Q9XHT1 brassica ol
33	22	100.0	72	17	Q97114	Q97114 sulfolobus
34	22	100.0	73	16	Q88WQ1	Q88WQ1 lactobacilli
35	22	100.0	76	5	Q8T6A2	Q8T6A2 aplysia cal
36	22	100.0	76	15	Q89369	Q89369 human immun
37	22	100.0	76	16	Q8RHG0	Q8RHG0 fusobacteri
38	22	100.0	77	2	Q8B142	Q8B142 vibrio para
39	22	100.0	77	10	Q84JQ2	Q84JQ2 arabidopsis
40	22	100.0	78	5	Q86W41	Q86W41 phylioperth
41	22	100.0	78	11	Q91Y28	Q91Y28 mus musculu
42	22	100.0	81	15	Q68996	Q68996 human endog
43	22	100.0	81	16	Q7UPR5	Q7UPR5 rhodospirill
44	22	100.0	82	16	Q8G4U1	Q8G4U1 bifidobacte
45	22	100.0	83	11	Q62306	Q62306 mus musculu

ALIGNMENTS

RESULT 1

Q94199 ID Q94199 PRELIMINARY; PRT; 25 AA.
AC Q94199; 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 40S ribosomal protein (Fragment).
OS Atropa belladonna (Belladonna) (Deadly nightshade).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Atropa.
OX NCBI_TaxID=33113;
RN [1]
RP SEQUENCE FROM N.A.
RA Nouar E., Baucher M., Jaziri M.;
RT "Differential gene expression in Atropa belladonna leafy gall.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA El Hassan N.;
RL Thesis (2001), Department of Plant Biotechnology,
RL Universite Libre De Bruxelles, Bruxelles, Belgium.
DR EMBL: A1921783; CAC4248.1; --
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0005840; C:ribosome; IEA.
DR GO: GO:0003735; F:structural constituent of ribosome; IEA.
DR GO: GO:006412; F:protein biosynthesis; IEA.
DR InterPro: IPR000754; Ribosomal_S9.
DR Pfam: PF00380; Ribosomal_S9; 1.
DR Pfam: PF001627; Ribosomal_S9; 1.
KW Ribosomal protein.
FT NON TER 1
SQ SEQUENCE 25 AA; 2986 MW; BFF06F7182304F69 CRC64;

Query Match 100.0%; Score 22; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
 DB 9 KKF 12

RESULT 2

O98MC2 PRELIMINARY; PRT; 30 AA.
 AC O98MC2;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein ms10637.
 OS Mesorhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_Taxid=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneke T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idehara K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP002995; BAB48191.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 30 AA; 3302 MW; 9DDIDA28C11C6577 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 30;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
 DB 2 KKF 5

RESULT 3
 ID O97875 PRELIMINARY; PRT; 35 AA.
 AC O97875;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Gp160 protein (Fragment).
 GN ENV OR ENV. V3.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VI526;
 RX MEDLINE=96426454; PubMed=8828748;
 RA Delaporte E., Janssens W., Peeters M., Buve A., D'Anga C.,
 RA Perret J.L., Ditsambou V., Georges Courtbot M.C., Georges A.,
 RA Bougeois A., Samb B., Henzel D., Heyndrickx L., Franssen K.,
 RA Van der Groen G., Larouz B., Mde J.R.;
 RT "Epidemiological and Molecular characteristics of HIV infection in
 RT Gabon (1986 - 1994)."
 RL AIDS 10:903-910(1996).
 DR EMBL: X90922; CAA62429.1;
 DR GO: GO:0019028; C:Viral capsid; IEA.
 DR GO: GO:0019031; C:viral envelope; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 KW AIDS; Coat protein; Glycoprotein.

FT NON TER 1 1
 FT NON TER 35 35
 SQ SEQUENCE 35 AA; 3847 MW; 81AF1F7D04503244 CRC64;

Query Match 100.0%; Score 22; DB 15; Length 35;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
 DB 12 KKF 15

RESULT 4

O97732 PRELIMINARY; PRT; 50 AA.
 AC O97732;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein SP0124.
 GN SP0124.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_Taxid=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC BAA-334 / TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapfe E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 DR EMBL: AE007329; AAK74310.1; -.
 DR PIR: E95014; E95014.
 DR TIGR: SP0124; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 50 AA; 5558 MW; 9393EB8E261530BF CRC64;

Query Match 100.0%; Score 22; DB 16; Length 50;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKF 4
 DB 44 KKF 47

RESULT 5

O9DRH5 PRELIMINARY; PRT; 50 AA.
 AC O9DRH5;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Orf51.
 GN ORF51 OR SP0127.
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_Taxid=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 HA Hopkins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,

RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA McAhren S.M., McHenry M., McMaster K., Mundy C.W., Nicot T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun F.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaekunas S.R., Kosteck P.R., Jr., Skatrud P.L.,
 RA Glass J.I.,
 RT "Genome of the bacterium *Streptococcus pneumoniae* strain R6.",
 DR EMBL: AEO08396; AAK98931.1; --
 DR PIR: G97887; G97887.
 KW Complete proteome.
 SQ SEQUENCE 50 AA; 5557 MW; 7E2888EE26153897 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 50;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
 Db 44 KKF 47

RESULT 6

Q9KFA9 PRELIMINARY; PRT; 58 AA.
 AC Q9KFA9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypochemical protein BH0576.
 GN BH0576.
 OS *Bacillus halodurans*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=8665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
 RT *halodurans* and genomic sequence comparison with *Bacillus subtilis*."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001509; BAB04295.1; --
 DR PIR: H83721; H83721.
 KW Hypochemical protein; Complete proteome.
 SQ SEQUENCE 58 AA; 6353 MW; 40D50F0E76BD5E3 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 58;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
 Db 30 KKF 33

RESULT 7

Q963B0 PRELIMINARY; PRT; 59 AA.
 AC Q963B0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Antiheliotic peptide cecropin A2.
 GN CECRA2.
 OS *Aedes albopictus* (Forest day mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
 OX NCBI_TaxID=7160;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Sun D., Fallon A.;
 RT "Characterization of genomic DNA encoding mosquito cecropins."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF394745; AAK81850.1; --
 SQ SEQUENCE 59 AA; 6183 MW; C5C1DFE3D8A8673 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 59;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
 Db 27 KKF 30

RESULT 8

Q86PR6 PRELIMINARY; PRT; 59 AA.
 AC Q86PR6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cecropin A.
 OS *Culex pipiens pipiens* (Northern house mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culex.
 OX NCBI_TaxID=38569;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Iowa State;
 RA Bartholomay L.C., Farid H.A., Ramzy R.M., Christensen B.M.;
 RT "Immune immunity in the *Culex pipiens*-*Wuchereria bancrofti* host-
 RT parasite relationship."
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY189808; AAO38516.1; --
 SQ SEQUENCE 59 AA; 6285 MW; 26CE3D7E823D9296 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 59;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKF 4
 Db 27 KKF 30

RESULT 9

Q831G6 PRELIMINARY; PRT; 59 AA.
 AC Q831G6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Transcriptional regulator, putative.
 GN EF2543.
 OS *Enterococcus faecalis* (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RX MEDLINE=22550857; PubMed=12663927;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
 RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
 RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Kouri H.,
 RA Urdack T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
 RT "Role of mobile DNA in the evolution of vancomycin-resistant
 RT *Enterococcus faecalis*."
 RL Science 299:2071-2074(2003).

DR EMBL; AE016955; AAC02256.1; --
 DR TIGR; EF2543; --
 KM Complete proteome.
 SQ SEQUENCE 59 AA; 6731 MW; FCCA51EF3BCAD24 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 59;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
 ||||
 DB 52 KKEG 55

RESULT 10
 035588 PRELIMINARY; PRT; 60 AA.

AC 035588;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE C11 protein (Fragment).
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OC NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Baby kidney;
 RX MEDLINE=95082951; PubMed=7990965;
 RA Prolova L., Le Goff X., Raemussen H.H., Chaperagin S., Druegon G.,
 RA Kress M., Arman I., Haenni A.L., Celis J.E., Philippe M., Justesen K.,
 RA Kisseliev L.;
 RT "A highly conserved eukaryotic protein family possessing properties of
 RT polypeptide chain release factor.";
 RL Nature 372:701-703(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Baby kidney;
 RX MEDLINE=97338483; PubMed=9195043;
 RA Ueberro B., Burwilaichitr L., Stansfield I., Taasan J.P., Le Goff X.,
 RA Kress M., Tuite M.F.;
 RT "Expression of the release factor eRF1 (Sup45p) gene of higher
 RT eukaryotes in yeast and mammalian tissues.";
 RL Biochimie 79:27-36(1997).
 DR EMBL; X81624; CAAS7280.1; --
 DR HSSP; P46055; 1DT9.
 DR InterPro; IPR005142; eRF1_3.
 DR Pfam; PF03465; eRF1_3; 1.
 FT NON TER 1
 SQ SEQUENCE 60 AA; 6785 MW; 57F075EA624FEF6 CRC64;

Query Match 100.0%; Score 22; DB 11; Length 60;
 Best Local Similarity 100.0%; Pred. No. 6.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
 ||||
 DB 6 KKEG 9

RESULT 11
 082207 PRELIMINARY; PRT; 60 AA.

AC 082207;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein PA0507.
 GN PA0507.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;

OC Thermoproteaceae; Pyrobaculum.
 OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX MEDLINE=21664397; PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 RT aerophilum.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL; AE009771; AAL62834.1; --
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 60 AA; 7290 MW; 89171FD9473B8AD2 CRC64;

Query Match 100.0%; Score 22; DB 17; Length 60;
 Best Local Similarity 100.0%; Pred. No. 6.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
 ||||
 DB 20 KKEG 23

RESULT 12
 08U4W9 PRELIMINARY; PRT; 61 AA.

AC 08U4W9;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE AGR_L_2351P.
 GN AGR_L_2351.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 OC NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Askew M., Halling C., Mullin L.,
 RA Houmel K., Gordon J., Vaudin D., Tarchouk O., Epp A., Liu F.,
 RA Woliam C., Allinger M., Douglas D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Guxon J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE008318; AAK89742.1; --
 DR PIR; D98277; D98277.
 SQ SEQUENCE 61 AA; 6772 MW; 61SDA6FA3A0A56D2 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 61;
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKEG 4
 ||||
 DB 32 KKEG 35

RESULT 13
 07VCU4 PRELIMINARY; PRT; 61 AA.

AC 07VCU4;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Predicted protein family PM-6.
 GN PRO0646.
 OS Prochlorococcus marinus.
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.
 OX NCBI_TaxID=1219;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SARG / CCMP 1375 / SS120;
 RX MEDLINE=22810154; PubMed=12917486;
 RA Diferene A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
 Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
 RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
 RA Scanlan D.J., Tandeau de Marsac N., Weisenbach J., Wincker P.,
 RA Wolf Y.I., Hess W.R.;
 RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
 RT a nearly minimal oxygenotrophic genome.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
 DR EMBL; AB017162; AAP9690.1; -.
 KM Complete proteome.
 SQ SEQUENCE 61 AA; 6747 MW; AC032EF5BF7997B CRC64;

Query Match 100.0%; Score 22; DB 16; Length 61;
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXFG 4
 DB 12 KXFG 15

RESULT 14

O9LOG1 PRELIMINARY; PRT; 62 AA.
 AC O9LOG1;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
 DE P1504.24.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Becker J.R.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Becker J.R.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooke S., Buehler E., Chao O., Johnson-Hopson C.,
 RA Khan S., Kim C., Altfati H., Bei B., Chin C., Chien J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharzky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Becker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007887; AAF79352.1; -.
 SQ SEQUENCE 62 AA; 7607 MW; DD8DF41F1624FEC4 CRC64;

Query Match 100.0%; Score 22; DB 10; Length 62;
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXFG 4
 DB 24 KXFG 27

RESULT 15
 P76618 PRELIMINARY; PRT; 62 AA.
 ID P76618

AC P76618;
 DT 01-FEB-1997 (TEMBLrel. 02, Created)
 DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE From BARS 275730 to 2786129 (Section 240 of 400) of the complete
 DE genome (Section 240 of 400).
 GN B2656.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.U.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 DR EMBL; AE000350; AAC75703.1; -.
 DR PIR; A65045; A65045.
 KM Complete proteome.
 SQ SEQUENCE 62 AA; 6714 MW; ED2C3CC9D87C20A2 CRC64;

Query Match 100.0%; Score 22; DB 16; Length 62;
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXFG 4
 DB 15 KXFG 18

Search completed: March 18, 2004, 11:19:19
 Job time : 3.09409 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:53:24 ; Search time 4.9628 Seconds
(without alignments)
683.197 Million cell updates/sec

Title: US-10-076-421-4
Perfect score: 67
Sequence: 1 ASTDTWGRPLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	100.0	12	5	AAG79462 sc-uPA 10
2	67	100.0	86	7	ABR42617 Human abr
3	67	100.0	86	7	ABR42602 Mouse abr
4	67	100.0	86	7	ABR42599 Human abr
5	67	100.0	86	7	ABR42601 Human abr
6	67	100.0	87	7	ABR42604 Human abr
7	67	100.0	87	7	ABR42605 Human abr
8	67	100.0	88	5	AAE16542 Human uro
9	67	100.0	89	2	AAW22742 Urokinase
10	67	100.0	96	5	AAE16550 Human uPA
11	67	100.0	100	1	AAE80996 Engineere
12	67	100.0	135	5	AAE16545 Human uro
13	67	100.0	143	5	AAE16549 Human uPA
14	67	100.0	201	2	AAW22746 Metastasi
15	67	100.0	208	2	AAW22747 Metastasi
16	67	100.0	322	7	ABR42615 Fusion pr
17	67	100.0	322	7	ABR42616 Fusion pr
18	67	100.0	337	4	AAG79462 Human col
19	67	100.0	337	5	ABP41795 Human ova
20	67	100.0	365	2	AAE68854 Delta 1-4
21	67	100.0	378	2	AAW13635 Human pro
22	67	100.0	386	2	AAE66266 Bifunctio
23	67	100.0	389	2	AAW13636 Human pro
24	67	100.0	390	2	AAE66245 Bifunctio
25	67	100.0	390	2	AAE66247 Bifunctio

26	67	100.0	392	2	AAE66260 Bifunctio
27	67	100.0	392	2	AAE66264 Bifunctio
28	67	100.0	392	2	AAE66255 Bifunctio
29	67	100.0	392	2	AAE66259 Bifunctio
30	67	100.0	392	2	AAE66261 Bifunctio
31	67	100.0	392	2	AAE66258 Bifunctio
32	67	100.0	392	2	AAE66263 Bifunctio
33	67	100.0	392	2	AAE66254 Bifunctio
34	67	100.0	392	2	AAE66256 Bifunctio
35	67	100.0	392	2	AAE66257 Bifunctio
36	67	100.0	393	2	AAE66251 Bifunctio
37	67	100.0	393	2	AAE66253 Bifunctio
38	67	100.0	393	2	AAE66249 Bifunctio
39	67	100.0	393	2	AAE66252 Bifunctio
40	67	100.0	393	2	AAE66244 Bifunctio
41	67	100.0	393	2	AAE66250 Bifunctio
42	67	100.0	393	2	AAE66257 Bifunctio
43	67	100.0	393	2	AAE66256 Bifunctio
44	67	100.0	393	2	AAE66257 Bifunctio
45	67	100.0	395	2	AAE66247 Bifunctio

ALIGNMENTS

RESULT 1
AAG79462
ID AAG79462 standard; peptide; 12 AA.
XX
AC AAG79462:
XX
DT 15-NOV-2002 (first entry)
XX
DE sc-uPA long A chain peptide fragment 2, amino acids 82-93.
XX
KM Single chain prepro-urokinase; sc-uPA; pro-urokinase; HIV;
KM high molecular weight urokinase-type plasminogen activator; HMW-uPA;
KM long A; long B; EGF-like domain; kringle domain; urokinase receptor;
KM low molecular weight urokinase-type plasminogen activator; LMW-uPA; CD87;
KM binding domain.
XX
OS Homo sapiens.
XX
PN EP1232755-A2.
XX
PD 21-AUG-2002..
XX
PF 15-FEB-2002; 2002EP-00003555.
XX
PR 20-FEB-2001; 2001UP-00042655.
PR 19-JUN-2001; 2001UP-00184284.
XX
PA (JGRP-) JCR PHARM CO LTD.
XX
PI Wada M, Wada N;
XX
DR WPI; 2002-610512/66.
XX
PT Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g. high
PT molecular weight urokinase-type plasminogen activator, amino-terminal
PT fragment or an anti-CD87 antibody.
XX
PS Example; Page 23; 38pp; English.
XX
CC The sequences given in AAG79461-63 represent peptide fragments derived
CC from the long A chain of single chain prepro-urokinase (sc-uPA). Pro-
CC urokinase (amino acids 21-431) with a cleavage between amino acids 178
CC and 179 gives high molecular weight urokinase-type plasminogen activator
CC (HMW-uPA). HMW-uPA is a protein consisting of two peptide chains linked
CC by a disulphide bond. The chains, long A and B, are formed by enzymatic
CC cleavage between amino acids 178 and 179 of pro-urokinase. HMW-uPA
CC includes an EGF-like domain, a kringle domain and a urokinase receptor
CC (CD87) binding domain. HMW-uPA is then cleaved between amino acids 155

CC and 156 to give low molecular weight urokinase-type plasminogen activator (LMW-uPA) (amino acids 156-178 and 179-431), that has no plasminogen activator activity. SC-uPA, or these fragments of it, may be used in the anti-HIV agents of the invention which comprise a ligand molecule that binds to CD87. The agents are useful for treating HIV-infected humans for suppression of reproduction of HIV. The anti-HIV agents act by a mechanism of action different from those of conventional drugs, widening the choice of therapeutics agents and avoiding problems of resistant HIV

CC
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 1 ASTDTMGRPCLP 12

RESULT 2
ABR42617
ID ABR42617 standard; protein; 86 AA.
XX
AC ABR42617;
XX
DT 26-AUG-2003 (first entry)
XX
DE Human abrogen (hATF-kringle).
XX
KW Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KM urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 43
FT /label= Asn, Asp
FT Misc-difference 74
FT /label= Pro, Leu
XX
PN WO2003042354-A2.
XX
PD 22-MAY-2003.
XX
PF 04-SEP-2002; 2002WO-US027885.
XX
PR 04-SEP-2001; 2001US-0316300P.
XX
PA (AVET) AVENTIS PHARM INC.
XX
PI Nesbitt M, Fong TC, Brockstedt D;
XX
DR WPI; 2003-449566/42.
DR N-PSDB; ACC58338.
XX
PT New abrogen polypeptide, useful for treating an angiogenesis related
FT diseases e.g. tumor metastasis.
XX
PS Disclosure; Page 95; 95pp; English.
XX
CC The present sequence is the protein sequence of a novel human abrogen,
CC designated hATF-kringle, comprising the human urokinase plasminogen
CC activator kringle domain. Abrogens such as hATF-kringle are potent
CC inhibitors of endothelial proliferation and angiogenesis. Abrogen
CC polypeptides are capable of inhibiting or reducing cell proliferation
CC induced by both basic fibroblast growth factor (bFGF) and vascular
CC endothelial growth factor in a specific endothelial cell proliferation
CC assay; angiotensin only inhibits bFGF induced proliferation in this
CC assay. Vectors that expressed abrogen polypeptides in vivo were shown to
CC reduce tumour metastasis in 2 lung cancer models. The invention provides
CC abrogen polypeptides and polynucleotides, and methods of using these to

CC treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
CC (claimed)
XX
SQ Sequence 86 AA;

Query Match 100.0%; Score 67; DB 7; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.00073;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 15 ASTDTMGRPCLP 26

RESULT 3
ABR42602
ID ABR42602 standard; protein; 86 AA.
XX
AC ABR42602;
XX
DT 26-AUG-2003 (first entry)
XX
DE Mouse abrogen (hATF-kringle).
XX
KW Mouse; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KM urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW gene therapy.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Domain 53..59
FT /note= "Xkringle domain"
XX
PN WO2003042354-A2.
XX
PD 22-MAY-2003.
XX
PF 04-SEP-2002; 2002WO-US027885.
XX
PR 04-SEP-2001; 2001US-0316300P.
XX
PA (AVET) AVENTIS PHARM INC.
XX
PI Nesbitt M, Fong TC, Brockstedt D;
XX
DR WPI; 2003-449566/42.
DR N-PSDB; ACC58337.
XX
PT New abrogen polypeptide, useful for treating an angiogenesis related
FT diseases e.g. tumor metastasis.
XX
PS Claim 1; Page 25; 95pp; English.
XX
CC The present sequence is the protein sequence of a novel mouse abrogen,
CC designated hATF-kringle, comprising the human urokinase plasminogen
CC activator kringle domain. Abrogens such as hATF-kringle are potent
CC inhibitors of endothelial proliferation and angiogenesis. Abrogen
CC polypeptides are capable of inhibiting or reducing cell proliferation
CC induced by both basic fibroblast growth factor (bFGF) and vascular
CC endothelial growth factor in a specific endothelial cell proliferation
CC assay; angiotensin only inhibits bFGF induced proliferation in this
CC assay. Vectors that expressed abrogen polypeptides in vivo were shown to
CC reduce tumour metastasis in 2 lung cancer models. The invention provides
CC abrogen polypeptides and polynucleotides, and methods of using these to
CC treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
CC (claimed)
XX
SQ Sequence 86 AA;

Query Match 100.0%; Score 67; DB 7; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.00073;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
 |||||
 Db 15 ASTDTMGRPCLP 26

RESULT 4

ABR42599
 ID ABR42599 standard; protein; 86 AA.
 XX
 AC ABR42599;

DT 26-AUG-2003 (first entry)

DE Human abrogen (hATF-kringle).

XX Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
 KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
 KM gene therapy.

XX Homo sapiens.

FT Key Location/Qualifiers
 Domain 53..59
 /note="kringle domain"

XX WO2003042354-A2.

PD 22-MAY-2003.

PF 04-SEP-2002; 2002WO-US027885.

PR 04-SEP-2001; 2001US-0316300P.

XX (AVET) AVENTIS PHARM INC.

XX Nesbit M, Fong TC, Brockstedt D;

XX WPI; 2003-449566/42.

DR N-PSDB; ACCS8334.

PT New abrogen polypeptide, useful for treating an angiogenesis related
 diseases e.g. tumor metastasis.

PS Claim 1; Page 24; 95pp; English.

XX The present sequence is the protein sequence of a novel human abrogen,
 CC designated hATF-kringle, comprising the human urokinase plasminogen
 CC activator kringle domain. Abrogens such as hATF-kringle are potent
 CC inhibitors of endothelial proliferation and angiogenesis. Abrogen
 CC polypeptides are capable of inhibiting or reducing cell proliferation
 CC induced by both basic fibroblast growth factor (bFGF) and vascular
 CC endothelial growth factor in a specific endothelial cell proliferation
 CC assay; angiostatin only inhibits bFGF induced proliferation in this
 CC assay. Vectors that expressed abrogen polypeptides in vivo were shown to
 CC reduce tumour metastasis in 2 lung cancer models. The invention provides
 CC abrogen polypeptides and polynucleotides, and methods of using these to
 CC treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
 CC (claimed)

XX Sequence 86 AA;

QY Query Match 100.0%; Score 67; DB 7; Length 86;
 Best Local Similarity 100.0%; Pred. No. 0.00073;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
 |||||
 Db 15 ASTDTMGRPCLP 26

RESULT 5

ABR42601

ID ABR42601 standard; protein; 86 AA.

XX ABR42601;

DT 26-AUG-2003 (first entry)

DE Human abrogen (hATF-kringle).

XX Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
 KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
 KM gene therapy.

XX Homo sapiens.

FT Key Location/Qualifiers
 Domain 53..59
 /note="kringle domain"

XX WO2003042354-A2.

PD 22-MAY-2003.

PF 04-SEP-2002; 2002WO-US027885.

PR 04-SEP-2001; 2001US-0316300P.

XX (AVET) AVENTIS PHARM INC.

XX Nesbit M, Fong TC, Brockstedt D;

XX WPI; 2003-449566/42.

DR N-PSDB; ACCS8336.

PT New abrogen polypeptide, useful for treating an angiogenesis related
 diseases e.g. tumor metastasis.

PS Claim 1; Page 25; 95pp; English.

XX The present sequence is the protein sequence of a novel human abrogen,
 CC designated hATF-kringle, comprising the human urokinase plasminogen
 CC activator kringle domain. Abrogens such as hATF-kringle are potent
 CC inhibitors of endothelial proliferation and angiogenesis. Abrogen
 CC polypeptides are capable of inhibiting or reducing cell proliferation
 CC induced by both basic fibroblast growth factor (bFGF) and vascular
 CC endothelial growth factor in a specific endothelial cell proliferation
 CC assay; angiostatin only inhibits bFGF induced proliferation in this
 CC assay. Vectors that expressed abrogen polypeptides in vivo were shown to
 CC reduce tumour metastasis in 2 lung cancer models. The invention provides
 CC abrogen polypeptides and polynucleotides, and methods of using these to
 CC treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
 CC (claimed)

XX Sequence 86 AA;

QY Query Match 100.0%; Score 67; DB 7; Length 86;
 Best Local Similarity 100.0%; Pred. No. 0.00073;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
 |||||
 Db 15 ASTDTMGRPCLP 26

RESULT 6

ABR42604
 ID ABR42604 standard; protein; 87 AA.

XX ABR42604;

DT 26-AUG-2003 (first entry)

DE Human abrogen (hATF-kringle).

KW	Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KM	urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW	gene therapy.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Domain 54..60
FT	/note= "kringle domain"
PD	WO2003042354-A2.
XX	
PD	22-MAY-2003.
XX	
PF	04-SEP-2002; 2002MO-US027885.
XX	
PR	04-SEP-2001; 2001US-0316300P.
XX	
PA	(AVET) AVENTIS PHARM INC.
XX	
PI	Nesbit M, Fong TC, Brockstedt D;
XX	
DR	WPI; 2003-449566/42.
XX	
PT	New abrogen polypeptide, useful for treating angiogenesis related
PT	diseases e.g. tumor metastasis.
XX	
PS	Claim 37; Page 26; 95dp; English.
XX	
CC	The present sequence is the protein sequence of a secreted form of a
CC	novel human abrogen, designated hATF-kringle, comprising the human
CC	urokinase plasminogen activator kringle domain. The sequence includes an
CC	N-terminal alanine residue that results from cleavage of an interleukin-2
CC	signal peptide. Abrogens such as hATF-kringle are potent inhibitors of
CC	endothelial proliferation and angiogenesis. Abrogen polypeptides are
CC	capable of inhibiting or reducing cell proliferation induced by both
CC	basic fibroblast growth factor (bFGF) and vascular endothelial growth
CC	factor in a specific endothelial cell proliferation assay; angiosatin
CC	only inhibits bFGF induced proliferation in this assay. Vectors that
CC	expressed abrogen polypeptides in vivo were shown to reduce tumour
CC	metastasis in 2 lung cancer models. The invention provides abrogen
CC	polypeptides and polymucleotides, and methods of using these to treat an
CC	angiogenesis-related disease or disorder, e.g. tumour metastasis
CC	(claimed)
XX	
SQ	Sequence 87 AA:
	Query Match 100.0%; Score 67; DB 7; Length 87;
	Best Local Similarity 100.0%; Pred. No. 0.00074;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY	1 ASTDTMGRCPLP 12
DB	16 ASTDTMGRCPLP 27
RESULT 7	
ABR42605	
ID	ABR42605 standard; protein; 87 AA.
XX	
AC	ABR42605;
XX	
DT	26-AUG-2003 (first entry)
XX	
DE	Human abrogen (hATF-kringle).
XX	
KW	Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KM	urokinase plasminogen activator; tumour; metastasis; cyostatic;
KW	gene therapy.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers

FT	Domain	54..60	/note="kringle domain"
XX			
PN	WO2003042354-A2.		
PD	22-MAY-2003.		
XX			
PF	04-SEP-2002; 2002WO-US027885.		
XX			
PR	04-SEP-2001; 2001US-0316300P.		
XX			
PA	(AVET) AVENTIS PHARM INC.		
XX			
PI	Neablt M, Fong TC, Brockstedt D;		
XX			
DR	WPI; 2003-449566/42.		
XX			
PT	New abrogen polypeptide, useful for treating an angiogenesis related		
XX	diseases e.g. tumor metastasis.		
PS	Claim 37; Page 26; 95pp; English.		
XX			
CC	The present sequence is the protein sequence of a secreted form of a		
CC	novel human abrogen, designated hATP-kringle, comprising the human		
CC	urokinase plasminogen activator kringle domain. The sequence includes an		
CC	N-terminal alanine residue that results from cleavage of an interleukin-2		
CC	signal peptide. Abrogens such as hATP-kringle are potent inhibitors of		
CC	endothelial proliferation and angiogenesis. Abrogen polypeptides are		
CC	capable of inhibiting or reducing cell proliferation induced by both		
CC	basic fibroblast growth factor (bFGF) and vascular endothelial growth		
CC	factor in a specific endothelial cell proliferation assay; angiotatin		
CC	only inhibits bFGF induced proliferation in this assay. Vectors that		
CC	expressed abrogen polypeptides in vivo were shown to reduce tumour		
CC	metastasis in 2 lung cancer models. The invention provides abrogen		
CC	polypeptides and polynucleotides, and methods of using these to treat an		
CC	angiogenesis-related disease or disorder, e.g. tumour metastasis		
CC	(claimed)		
XX			
SQ	Sequence 87 AA;		
	Query Match	100.0%;	Score 67; DB 7; Length 87;
	Best Local Similarity	100.0%;	Pred. No. 0.00074;
	Matches 12; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 ASTDTMGRPCLP 12		
DB	16 ASTDTMGRPCLP 27		
RESULT 8			
AAE16542			
ID	AAE16542 standard; protein; 88 AA.		
XX			
AC	AAE16542;		
XX			
DT	09-APR-2002 (first entry)		
DE			
XX	Human urokinase-type plasminogen activator (uPA) kringle.		
XX			
XX	Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;		
KW	stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;		
KW	microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;		
KW	tumour cell metastasis; glioma; diabetic retinopathy; wound healing;		
KW	clotting disorder; uterine contraction disorder; respiratory disease;		
KW	male impotence; adult respiratory distress syndrome.		
OS	Homo sapiens.		
XX			
PN	WO200197752-A2.		
XX			
PD	27-DEC-2001.		
XX			
PF	13-JUN-2001; 2001WO-US018976.		

XX 20-JUN-2000; 2000US-0212874P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 XX Cines DB, Higazi AA;
 XX WPI; 2002-122240/16.
 DR N-PSDB; AAD27075.
 XX
 PT Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator.
 XX
 PS Claim 1; Fig 1A; 117pp; English.
 XX
 CC The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle
 CC
 SQ Sequence 88 AA;

Query Match 100.0%; Score 67; DB 5; Length 88;
 Best Local Similarity 100.0%; Pred. No. 0.00075; Indels 0; Gaps 0;
 Matches 12; Conservative 0; Mismatches 0;

QY 1 ASTDTMGRPCLP 12
 |||||
 DB 15 ASTDTMGRPCLP 26

RESULT 9
 AAM22742
 ID AAM22742 standard; protein; 89 AA.
 XX
 AC AAM22742;

XX 12-MAR-1998 (first entry)
 XX
 DE Urokinase residues 43-131.
 XX

KM Chimeric; human; urinary trypsin; inhibitor; HI-8; cancer; metastasis;
 KM G-domain; urokinase; prevention; leukaemia; lymphoma.
 XX

OS Homo sapiens.
 XX

PN MO9725422-A1.
 XX

PD 17-JUL-1997.
 XX

PF 06-JAN-1997; 97MO-JP000008.
 XX

PR 08-JAN-1996; 96JP-00001059.
 XX

PA (NISP) NISSIN FOOD PROD CO LTD.
 XX

PI Kobayashi H, Terao T, Sugino D, Okushima M;
 XX

DR WPI; 1997-372862/34.
 XX

PT Chimeric protein which inhibits development of metastases in cancer -
 PT contains urinary trypsin inhibitor carboxy-terminal domain linked to

PT urokinase G-domain.
 XX
 PS Claim 3; Page 72; 97pp; Japanese.
 XX

CC A novel chimeric protein contains the carboxy-terminal domain of human
 CC urinary trypsin inhibitor (HI-8), which inhibits cancer cell metastasis,
 CC linked to a peptide containing the G-domain of urokinase (AAM22742),
 CC which specifically binds the excess urokinase receptor expressed in
 CC cancer cells. The chimeric protein has the amino-terminal AAM22734, the
 CC carboxy-terminal AAM22735 and a linking sequence selected from AAM22736-
 CC 39 or partial sequences derived from these, specifically AAM38130-63. The
 CC chimeric protein may also have additional amino-terminal sequences
 CC selected from AAM22740 or 9 partial sequences derived from this, and/or
 CC additional carboxy-terminal sequences selected from AAM22743 or 10
 CC partial sequences derived from this. The chimeric protein can be used to
 CC prevent metastasis in, e.g. cancer of the lung, kidney, pancreas,
 CC stomach, colon, rectum, ovary, uterus, brain, skin, muscle, breast or
 CC prostate, and in leukaemia or lymphoma
 CC
 SQ Sequence 89 AA;

Query Match 100.0%; Score 67; DB 2; Length 89;
 Best Local Similarity 100.0%; Pred. No. 0.00076; Indels 0; Gaps 0;
 Matches 12; Conservative 0; Mismatches 0;

QY 1 ASTDTMGRPCLP 12
 |||||
 DB 20 ASTDTMGRPCLP 31

RESULT 10
 AAE16550
 ID AAE16550 standard; protein; 96 AA.
 XX
 AC AAE16550;

XX 09-APR-2002 (first entry)
 XX

DE Human uPA kringle and connecting peptide.
 XX

KM Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KM stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KM microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KM tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KM clotting disorder; uterine contraction disorder; respiratory disease;
 KM adult respiratory distress syndrome; male impotence.
 XX

OS Homo sapiens.
 XX

PN MO200197752-A2.
 XX

PD 27-DEC-2001.
 XX

PF 13-JUN-2001; 2001WO-US018976.
 XX

PR 20-JUN-2000; 2000US-0212874P.
 XX

PA (UYPE-) UNIV PENNSYLVANIA.
 XX

PI Cines DB, Higazi AA;
 XX

DR WPI; 2002-122240/16.
 XX

DR N-PSDB; AAD27083.
 XX

PT Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator.
 XX

PS Claim 25; Fig 11; 117pp; English.
 XX

CC The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,

endothelial cell or tissue. The composition is used for treating stroke, hypotension, hypertension, atherosclerosis, heart attack, microvascular occlusions, thrombotic microangiopathies, surgically induced thrombotic disorders, angiotenic disorders, pulmonary fibrosis, asthma, tumour cell invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma, diabetic retinopathy, wound healing, clotting disorder, uterine contraction disorder, male impotence, respiratory disease or condition such as asthma, adult respiratory distress syndrome, primary pulmonary hypertension, microvascular thrombotic occlusion, and a disorder associated with chronic intrapulmonary fibrin formation. The present sequence is human urokinase-type plasminogen activator (uPA) kringle and connecting peptide

Sequence 96 AA;

Query Match 100.0%; Score 67; DB 5; Length 96;
 Best Local Similarity 100.0%; Pred. No. 0.00082;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ASTDTMGRPCLP 12
 15 ASTDTMGRPCLP 26

Db

RESULT 11

AA080996 ID A080996 standard; protein, 100 AA.

AA080996; AC A080996; AD 25-MAR-2003 (revised)
 DT 08-OCT-1990 (first entry)

DE Engineered Human prourokinase mutant.

XX human pro-urokinase mutant; fibrinolysis; epidermal growth factor domain;
 KW ss.

XX Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 70
 FT /label= mutation
 FT /note= "Wild-type= Cys; Mutant= Ser."

XX EP253241-A.
 XX 20-JAN-1988.
 XX 03-JUN-1987; 87EP-00109628.
 XX 03-JUN-1986; 86JP-00156936.
 XX 18-FEB-1987; 87JP-00036495.
 XX (GREC) GREEN CROSS CORP.

XX Kaesi S, Hiramatsu R, Uno S, Nagai M, Arimura H;
 PI WPI; 1988-015623/03.
 XX N-PSDB; AAN81299.

XX New human pro-urokinase mutants with fibrinolytic activity - have longer
 PT half-lives in blood and are obtained by recombinant DNA procedures.

XX Disclosure; Page 7; 40p; English.

XX The DNA encoding this peptide contains two mutations introduced by site -
 CC directed mutagenesis to create two unique restriction sites; the first is
 CC recognised by SacI and the second by NdeI. The mutant amino acid denoted
 CC in the features is encoded by the first 3 bases of the NdeI site. The
 CC invention covers three deletion mutants derived from pro-kinase. All
 CC three begin from amino acid 21 (Ser) and have either 1) Asn(30) to
 CC Glu(63) and Thr(69) to Ser(70), or 2) Asn(30) to Lys(66) and Ser(70), or

3) Asn(30) to Tyr(71) deleted. The deleted region contains the Epidermal
 CC Growth Factor domain of the enzyme. (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 100 AA;

Query Match 100.0%; Score 67; DB 1; Length 100;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ASTDTMGRPCLP 12
 82 ASTDTMGRPCLP 93

Db

RESULT 12

AAE16545 ID AAE16545 standard; protein, 135 AA.

AAE16545; AC AAE16545; AD 09-APR-2002 (first entry)

DE Human urokinase-type plasminogen activator amino terminal fragment (ATF).

XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiotenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW adult respiratory distress syndrome; amino terminal fragment; ATF;
 KW male impotence.

XX Homo sapiens.
 XX W020019752-A2.
 XX 27-DEC-2001.
 XX 13-JUN-2001; 2001WO-US018976.
 XX 20-JUN-2000; 2000US-0212874P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 XX Cines DB, Higazi AA;
 XX WPI; 2002-122240/16.
 XX N-PSDB; AAD27078.

XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 CC comprising domains from urokinase-type plasminogen activator.

XX Claim 11; Fig 1D; 117p; English.

XX The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiotenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) amino
 CC terminal fragment (ATF)

XX Sequence 135 AA;

Query Match 100.0%; Score 67; DB 5; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 62 ASTDTMGRPCLP 73

RESULT 13

AAE16549
ID AAE16549 standard; protein; 143 AA.

XX AAE16549;

XX 09-APR-2002 (first entry)

XX Human uPA amino terminal fragment (ATF) and connecting peptide.

XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KM stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KM microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KM tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KM clotting disorder; uterine contraction disorder; respiratory disease;
KM adult respiratory distress syndrome; amino terminal fragment; ATF;
KM male impotence.

XX Homo sapiens.

XX WO200197752-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US018976.

XX 20-JUN-2000; 2000US-0212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

XX N-PSDB; AAD27082.

XX Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator.

XX Claim 24; Fig 1H; 117pp; English.

XX The invention relates to a composition comprising one or more domains of
CC urokinase-type plasminogen activator (uPA). The composition is used to
CC modulate the contractility and angiogenic activity of a mammalian muscle,
CC endothelial cell or tissue. The composition is used for treating stroke,
CC hypertension, atherosclerosis, heart attack, microvascular
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic interstitial pulmonary fibrosis. The present
CC sequence is human urokinase-type plasminogen activator (uPA) amino
CC terminal fragment (ATF) and connecting peptide

XX Sequence 143 AA;

Query Match 100.0%; Score 67; DB 5; Length 143;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 62 ASTDTMGRPCLP 73

RESULT 14

AAW22746
ID AAW22746 standard; protein; 201 AA.

XX AAW22746;

XX 12-MAR-1998 (first entry)

XX Metastasis inhibitor.

XX Chimeric; human; urinary trypsin; inhibitor; HI-8; cancer; metastasis;
KM G-domain; urokinase; prevention; leukaemia; lymphoma.

XX Homo sapiens.

XX WO9725422-A1.

XX 17-JUL-1997.

XX 06-JAN-1997; 97WO-JP000008.

XX 08-JAN-1996; 96JP-00001059.

XX (NISP) NISSIN FOOD PROD CO LTD.

XX Kobayashi H, Terao T, Sugino D, Okushima M;

XX WPI; 1997-372862/34.

XX N-PSDB; AAT75154.

XX Chimeric protein which inhibits development of metastases in cancer -
PT contains urinary trypsin inhibitor carboxy-terminal domain linked to
PT urokinase G-domain.

XX Claim 18; Page 57-58; 97pp; Japanese.

XX The present sequence is a novel chimeric protein, which contains the
CC carboxy-terminal domain of human urinary trypsin inhibitor (HI-8), which
CC inhibits cancer cell metastasis, linked to a peptide containing the G-
CC domain of urokinase, which specifically binds the excess urokinase
CC receptor expressed in cancer cells. The chimeric protein can be used to
CC prevent metastasis in, e.g. cancer of the lung, kidney, pancreas,
CC stomach, colon, rectum, ovary, uterus, brain, skin, muscle, breast or
CC prostate, and in leukaemia or lymphoma

XX Sequence 201 AA;

Query Match 100.0%; Score 67; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 63 ASTDTMGRPCLP 74

RESULT 15

AAW22747
ID AAW22747 standard; protein; 208 AA.

XX AAW22747;

XX 12-MAR-1998 (first entry)

XX Metastasis inhibitor.

XX Chimeric; human; urinary trypsin; inhibitor; HI-8; cancer; metastasis;
KM G-domain; urokinase; prevention; leukaemia; lymphoma.

XX Homo sapiens.
 OS
 XX
 PN WO9725422-A1.
 XX
 PD 17-JUL-1997.
 XX
 PF 06-JAN-1997; 97WO-JP000008.
 XX
 PR 08-JAN-1996; 96JP-00001059.
 XX
 PA (NISP) NISSIN FOOD PROD CO LTD.
 XX
 PI Kobayashi H, Terao T, Sugino D, Okushima M;
 XX
 DR WPI; 1997-372862/34.
 DR N-PSDB; AAT75155.
 XX
 PT Chimeric protein which inhibits development of metastases in cancer -
 PT contains urinary trypsin inhibitor carboxy-terminal domain linked to
 PT urokinase G-domain.
 XX
 PS Claim 19; Page 59-60; 97pp; Japanese.
 XX
 CC The present sequence is a novel chimeric protein, which contains the
 CC carboxy-terminal domain of human urinary trypsin inhibitor (HI-8) which
 CC inhibits cancer cell metastasis, linked to a peptide containing the G-
 CC domain of urokinase, which specifically binds the excess urokinase
 CC receptor expressed in cancer cells. The chimeric protein can be used to
 CC prevent metastasis in, e.g. cancer of the lung, kidney, pancreas,
 CC stomach, colon, rectum, ovary, uterus, brain, skin, muscle, breast or
 CC prostate, and in leukemia or lymphoma
 CC
 SQ Sequence 208 AA;

Query Match 100.0%; Score 67; DB 2; Length 208;
 Best Local Similarity 100.0%; Pred. No. 0.0018;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASTDTWGRPCLP 12
 |||||
 Db 63 ASTDTWGRPCLP 74

Search completed: March 18, 2004, 11:16:23
 Job time : 5.9628 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:13:05 ; Search time 1.39168 Seconds
(without alignments)
445.152 Million cell updates/sec

Title: US-10-076-421-4

Perfect score: 67

Sequence: 1 ASTDTMGRPCLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	89	US-09-101-272G-62	Sequence 62, Appl
2	67	100.0	138	US-08-797-689-12	Sequence 12, Appl
3	67	100.0	138	US-09-984-186-12	Sequence 12, Appl
4	67	100.0	157	US-08-142-5908-25	Sequence 25, Appl
5	67	100.0	194	US-09-101-372G-80	Sequence 80, Appl
6	67	100.0	200	US-09-101-372G-73	Sequence 73, Appl
7	67	100.0	201	US-09-101-372G-96	Sequence 96, Appl
8	67	100.0	208	US-09-101-372G-98	Sequence 98, Appl
9	67	100.0	365	US-08-093-741-83	Sequence 83, Appl
10	67	100.0	365	US-08-720-012-83	Sequence 83, Appl
11	67	100.0	393	US-08-560-098A-44	Sequence 44, Appl
12	67	100.0	393	US-08-967-024C-24	Sequence 24, Appl
13	67	100.0	393	US-08-967-024C-25	Sequence 25, Appl
14	67	100.0	411	US-08-087-163-1	Sequence 1, Appl
15	67	100.0	411	US-08-286-7488-18	Sequence 18, Appl
16	67	100.0	411	US-08-153-799-18	Sequence 18, Appl
17	67	100.0	411	US-08-560-098A-48	Sequence 48, Appl
18	67	100.0	411	US-09-181-816-1	Sequence 1, Appl
19	67	100.0	411	US-09-403-736-2	Sequence 2, Appl
20	67	100.0	430	US-07-942-157A-3	Sequence 3, Appl
21	67	100.0	430	US-07-942-157A-3	Sequence 3, Appl
22	67	100.0	431	US-09-101-272G-1	Sequence 1, Appl
23	67	100.0	431	US-08-560-098A-47	Sequence 47, Appl
24	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
25	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
26	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
27	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
28	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
29	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
30	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
31	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
32	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
33	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
34	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
35	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
36	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
37	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
38	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
39	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
40	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
41	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
42	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
43	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
44	67	100.0	432	US-08-051-142-2	Sequence 2, Appl
45	67	100.0	432	US-08-051-142-2	Sequence 2, Appl

28	38	56.7	593	2	US-08-468-249A-21	Sequence 21, Appl
29	37	55.2	288	4	US-09-252-991A-29045	Sequence 29045, A
30	36	53.7	138	4	US-09-134-000C-4574	Sequence 4574, Ap
31	36	53.7	160	4	US-09-800-729-191	Sequence 191, App
32	36	53.7	541	4	US-08-311-731A-171	Sequence 171, App
33	36	53.7	556	3	US-08-687-590-24	Sequence 24, Appl
34	35.5	53.0	118	4	US-09-252-991A-32474	Sequence 32474, A
35	35	52.2	135	4	US-09-252-991A-27145	Sequence 27145, A
36	35	52.2	137	4	US-09-489-039A-8628	Sequence 8628, Ap
37	35	52.2	202	4	US-09-540-336-3054	Sequence 3054, Ap
38	35	52.2	399	4	US-09-252-991A-25295	Sequence 25295, A
39	35	52.2	433	3	US-08-691-563C-87	Sequence 87, Appl
40	35	52.2	433	4	US-09-374-766-87	Sequence 87, Appl
41	35	52.2	433	4	US-08-979-847B-81	Sequence 81, Appl
42	35	52.2	525	4	US-09-252-991A-20491	Sequence 20491, A
43	35	52.2	787	4	US-09-721-183-2	Sequence 2, Appl
44	35	52.2	787	4	US-09-721-137-2	Sequence 2, Appl
45	35	52.2	787	4	US-09-721-251-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-101-272G-62
Sequence 62, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101, 272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.1
SEQ ID NO 62
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc feature
OTHER INFORMATION: residues 43-131 of the ATP domain of uPA
US-09-101-272G-62

Query Match 100.0%; Score 67; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCLP 12
Db 20 ASTDTMGRPCLP 31

RESULT 2
US-08-797-689-12
Sequence 12, Application US/08797689
Patent No. 5876969
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Guitton, Jean-Dominique
APPLICANT: Jung, Gerard
APPLICANT: Yeh, Patricia
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSER: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville

STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentln)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-12

Query Match 100.0%; Score 67; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCLP 12
|||||
Db 65 ASTDTMGRPCLP 76

RESULT 3
US-09-984-186-12
Sequence 12, Application US/09984186
Patent No. 6686179
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3c43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentln)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 100.0%; Score 67; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCLP 12
|||||
Db 65 ASTDTMGRPCLP 76

RESULT 4
US-08-142-590B-25
Sequence 25, Application US/08142590B
Patent No. 6120765
GENERAL INFORMATION:
APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadahito; HORII, Izumi; and
GOETTINCK,
TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS: 25
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,590B
FILING DATE: 25-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,318
FILING DATE: 02-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-009CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-142-590B-25

Query Match 100.0%; Score 67; DB 3; Length 157;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
DB 62 ASTDTMGRPCLP 73

RESULT 5
US-09-101-272G-80
; Sequence 80, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 80
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI chimeric protein
US-09-101-272G-80

Query Match 100.0%; Score 67; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
DB 63 ASTDTMGRPCLP 74

RESULT 6
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATP domain of uPA
US-09-101-272G-73

Query Match 100.0%; Score 67; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||

DB 82 ASTDTMGRPCLP 93

RESULT 7
US-09-101-272G-96
; Sequence 96, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 96
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI-CL chimeric protein
US-09-101-272G-96

Query Match 100.0%; Score 67; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
DB 63 ASTDTMGRPCLP 74

RESULT 8
US-09-101-272G-98
; Sequence 98, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 98
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI-ML chimeric protein
US-09-101-272G-98

Query Match 100.0%; Score 67; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
|||||
DB 63 ASTDTMGRPCLP 74

RESULT 9
US-08-093-741-83
; Sequence 83, Application US/08093741
; Patent No. 5681721
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERD J.
; APPLICANT: WENNDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES

APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,741
FILING DATE: 20-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-093-741-83

Query Match 100.0%; Score 67; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRCLP 12
DB 16 ASTDTMGRCLP 27

RESULT 10
US-08-720-012-83
Sequence 83, Application US/08720012
Patent No. 5747291
GENERAL INFORMATION:
APPLICANT: STEFFENS, GERD J.
APPLICANT: WENENDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-720-012-83

Query Match 100.0%; Score 67; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRCLP 12
DB 16 ASTDTMGRCLP 27

RESULT 11
US-08-560-098A-44
Sequence 44, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENENDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-560-098A-44

Query Match 100.0%; Score 67; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPLP 12
Db 17 ASTDTMGRPLP 28

RESULT 12
US-08-967-024C-24
Sequence 24, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOSHA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 100.0%; Score 67; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPLP 12
Db 17 ASTDTMGRPLP 28

RESULT 13
US-08-967-024C-25
Sequence 25, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: STEFFENS, Gerd Josef

APPLICANT: JANOSHA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 100.0%; Score 67; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPLP 12
Db 17 ASTDTMGRPLP 28

RESULT 14
US-08-087-163-1
Sequence 1, Application US/08087163
Patent No. 5472692
GENERAL INFORMATION:
APPLICANT: Liu, Jian-Ning
APPLICANT: Gurewich, Victor
TITLE OF INVENTION: PRO-UKINASE MUTANTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,163
FILING DATE: 07/02/93
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Faase, J. Peter
 REGISTRATION NUMBER: 32,983
 REFERENCE/DOCKET NUMBER: 04547/013001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 411
 TYPE: amino acid
 STRANDEDNESS: N/A
 TOPOLOGY: N/A
 US-08-087-163-1

Query Match 100.0%; Score 67; DB 1; Length 411;
 Best Local Similarity 100.0%; Pred. No. 0.0007;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASTDTMGRPCLP 12
 |||||
 Db 62 ASTDTMGRPCLP 73

RESULT 15
 US-08-286-748B-18
 Sequence 18, Application US/08286748B
 Patent No. 5759542
 GENERAL INFORMATION:
 APPLICANT: Victor Gurewich
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
 OF DRUGS BY PLATELETS FOR THE TREATMENT OF
 TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 55SX
 OPERATING SYSTEM: MS-DOS (Version 5.0)
 SOFTWARE: WordPerfect (Version 5.1)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/286,748B
 FILING DATE: August 5, 1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: J. Peter Faase
 REGISTRATION NUMBER: 32,983
 REFERENCE/DOCKET NUMBER: 04547/013001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 411
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-286-748B-18
 Query Match 100.0%; Score 67; DB 1; Length 411;

Best Local Similarity 100.0%; Pred. No. 0.0007;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASTDTMGRPCLP 12
 |||||
 Db 62 ASTDTMGRPCLP 73

Search completed: March 18, 2004, 11:21:22
 Job time: 1.39168 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:19:25 ; Search time 3.57112 Seconds
(without alignments)
870.166 Million cell updates/sec

Title: US-10-076-421-4
Perfect score: 67
Sequence: 1 ASTDTMGRPLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	12	13	US-10-076-421-4
2	67	100.0	86	15	US-10-233-675A-1
3	67	100.0	86	15	US-10-233-675A-1
4	67	100.0	86	15	US-10-233-675A-1
5	67	100.0	86	15	US-10-233-675A-1
6	67	100.0	86	15	US-10-233-675A-1
7	67	100.0	86	15	US-10-233-675A-1
8	67	100.0	86	15	US-10-233-675A-1
9	67	100.0	86	15	US-10-233-675A-1
10	67	100.0	86	15	US-10-233-675A-1
11	67	100.0	86	15	US-10-233-675A-1
12	67	100.0	86	15	US-10-233-675A-1
13	67	100.0	86	15	US-10-233-675A-1
14	67	100.0	86	15	US-10-233-675A-1
15	67	100.0	86	15	US-10-233-675A-1

16	67	100.0	138	14	US-10-237-871-12	Sequence 12, Appl
17	67	100.0	138	14	US-10-237-871-12	Sequence 12, Appl
18	67	100.0	138	9	US-09-880-503-8	Sequence 8, Appl
19	67	100.0	332	15	US-10-233-675A-20	Sequence 20, Appl
20	67	100.0	332	15	US-10-233-675A-21	Sequence 21, Appl
21	67	100.0	337	14	US-10-106-698-6266	Sequence 6266, Ap
22	67	100.0	403	9	US-10-264-049-2927	Sequence 2927, Ap
23	67	100.0	411	15	US-09-880-503-6	Sequence 6, Appl
24	67	100.0	411	15	US-09-880-503-3	Sequence 3, Appl
25	67	100.0	411	15	US-10-407-821-2	Sequence 2, Appl
26	67	100.0	431	9	US-09-264-468B-1	Sequence 1, Appl
27	67	100.0	431	12	US-10-411-037-34	Sequence 34, Appl
28	67	100.0	431	13	US-10-076-421-2	Sequence 2, Appl
29	67	100.0	431	14	US-10-171-311-184	Sequence 184, App
30	67	100.0	431	14	US-10-193-656-4	Sequence 4, Appl
31	67	100.0	431	14	US-10-301-822-161	Sequence 161, App
32	67	100.0	431	14	US-10-247-671-149	Sequence 149, App
33	67	100.0	431	14	US-10-131-985-21	Sequence 21, Appl
34	67	100.0	431	15	US-10-295-027-414	Sequence 414, App
35	67	100.0	445	15	US-10-295-027-1275	Sequence 1275, Ap
36	67	100.0	445	15	US-10-360-101-266	Sequence 266, App
37	67	100.0	650	15	US-10-401-077-1	Sequence 1, Appl
38	67	100.0	672	15	US-10-233-675A-15	Sequence 15, Appl
39	67	100.0	674	15	US-10-233-675A-14	Sequence 14, Appl
40	67	100.0	687	15	US-10-233-675A-17	Sequence 17, Appl
41	67	100.0	688	15	US-10-233-675A-18	Sequence 18, Appl
42	67	100.0	689	15	US-10-233-675A-13	Sequence 13, Appl
43	51	76.1	86	15	US-10-233-675A-3	Sequence 3, Appl
44	51	76.1	86	15	US-10-233-675A-24	Sequence 24, Appl
45	41	61.2	96	10	US-09-764-872-437	Sequence 437, App

ALIGNMENTS

RESULT 1
US-10-076-421-4
; Sequence 4, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: NAKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-076-421-4

Query Match 100.0%; Score 67; DB 13; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPLP 12
DB 1 ASTDTMGRPLP 12

RESULT 2
US-10-233-675A-1
; Sequence 1, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Neebit, Mark

```
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
; US-10-233-675A-1

Query Match          100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
Db      15 ASTDTMGRPCLP 26

RESULT 3
US-10-233-675A-5
; Sequence 5, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
; US-10-233-675A-5

Query Match          100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
Db      15 ASTDTMGRPCLP 26

RESULT 4
US-10-233-675A-7
; Sequence 7, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
```

```
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
; US-10-233-675A-7

Query Match          100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
Db      15 ASTDTMGRPCLP 26

RESULT 5
US-10-233-675A-22
; Sequence 22, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 22
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fragment of human urokinase plasminogen activator
; US-10-233-675A-22

Query Match          100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
Db      15 ASTDTMGRPCLP 26

RESULT 6
US-10-233-675A-27
; Sequence 27, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
```

```

; SOFTWARE: Patentin version 3.1
; SEQ ID NO 27
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human urokinase plasminogen activator
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (43)-(43)
; OTHER INFORMATION: Polymorphism - Xaa = Asn or Asp
; NAME/KEY: MISC FEATURE
; LOCATION: (74)-(74)
; OTHER INFORMATION: Polymorphism - Xaa = Pro or Leu
US-10-233-675A-27

Query Match          100.0%; Score 67; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
        |||||
Db       15 ASTDTMGRPCLP 26

RESULT 7
US-10-233-675A-9
; Sequence 9, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Neabdic, Mark
; APPLICANT: Pong, Timothy
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 9
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived fusion protein
US-10-233-675A-9

Query Match          100.0%; Score 67; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
        |||||
Db       16 ASTDTMGRPCLP 27

RESULT 8
US-10-233-675A-10
; Sequence 10, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Neabdic, Mark
; APPLICANT: Pong, Timothy
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
```

```

; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived fusion protein
US-10-233-675A-10

Query Match          100.0%; Score 67; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
        |||||
Db       16 ASTDTMGRPCLP 27

RESULT 9
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1

Query Match          100.0%; Score 67; DB 9; Length 88;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASTDTMGRPCLP 12
        |||||
Db       15 ASTDTMGRPCLP 26

RESULT 10
US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-9
```

Query Match 100.0%; Score 67; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
15 ASTDTMGRPCLP 26

Db 15 ASTDTMGRPCLP 26

RESULT 11
US-09-880-503-4
Sequence 4, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880.503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 135
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-4

Query Match 100.0%; Score 67; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
62 ASTDTMGRPCLP 73

Db 62 ASTDTMGRPCLP 73

RESULT 12
US-09-984-186-12
Sequence 12, Application US/09984186
Patent No. US20020151011A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984.186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797.689

FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: S792006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 100.0%; Score 67; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
65 ASTDTMGRPCLP 76

Db 65 ASTDTMGRPCLP 76

RESULT 13
US-10-237-667-12
Sequence 12, Application US/10237667
Publication No. US20030022308A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237.667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797.689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12

Query Match 100.0%; Score 67; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 65 ASTDTMGRPCLP 76

RESULT 14
US-10-237-708-12
Sequence 12, Application US/10237708
Publication No. US20030036170A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3c43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: word 5.1 (Patentln)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 100.0%; Score 67; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCLP 12
Db 65 ASTDTMGRPCLP 76

RESULT 15
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3c43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: word 5.1 (Patentln)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 100.0%; Score 67; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ASTDTMGRPCLP	12
Db	65	ASTDTMGRPCLP	76

Search completed: March 18, 2004, 11:36:07
Job time : 3.82112 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:10:10 ; Search time 1.23414 Seconds
(without alignments)
935.309 Million cell updates/sec

Title: US-10-076-421-4

Perfect score: 67

Sequence: 1 ASTDTMGRPCLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR 78: *
2: pir1: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	431	1 UKHU	u-plasminogen acti
2	52	77.6	433	1 UKBAY	u-plasminogen acti
3	51	76.1	432	1 S18932	u-plasminogen acti
4	51	76.1	443	1 UKMS	u-plasminogen acti
5	51	76.1	442	1 UKPG	u-plasminogen acti
6	40	59.7	224	2 S20463	siderophore biosyn
7	40	59.7	500	1 JCA153	cytochrome P450 2D
8	39	58.2	27	2 S29211	probable phosphol
9	39	58.2	288	2 T38696	probable peroxisom
10	39	58.2	419	2 S42989	T48 protein - frui
11	39	58.2	433	1 JN0560	u-plasminogen acti
12	39	58.2	543	2 AF1927	hypothetical prote
13	39	58.2	618	2 A35827	thrombin (EC 3.4.2
14	38	56.7	125	2 T49092	SB-like protein -
15	38	56.7	167	2 H72579	hypothetical prote
16	38	56.7	437	2 H86683	prophage p11 prote
17	38	56.7	589	2 I59297	parathyroid hormon
18	38	56.7	591	2 S44203	parathyroid hormon
19	38	56.7	591	2 I54195	parathyroid hormon
20	38	56.7	593	2 A49191	parathyroid hormon
21	37	55.2	153	2 B46353	ORF2 protein - coc
22	37	55.2	409	1 B0AG55	virB10 protein - A
23	37	55.2	773	2 A83888	hypothetical prote
24	37	55.2	804	2 T44606	transducer protein
25	37	55.2	810	2 T46810	halobacterial tran
26	37	55.2	810	2 P84327	HRF5 transducer [
27	37	55.2	1299	1 WMBEHS	membrane antigen p
28	37	55.2	1348	2 S27812	probable epidermal
29	37	55.2	1348	2 A43917	probable epidermal

30	36	53.7	231	2 H70968	hypothetical prote
31	36	53.7	247	2 T37820	hypothetical prote
32	36	53.7	322	2 E89057	protein K09H1.7 l
33	36	53.7	335	2 T31712	hypothetical prote
34	36	53.7	349	2 T15082	hypothetical prote
35	36	53.7	353	2 S05297	regulatory protein
36	36	53.7	374	2 F86243	ZIP4, probable zin
37	36	53.7	396	2 A82532	translation elonga
38	36	53.7	438	2 I50517	retinoid X recepto
39	36	53.7	529	2 B24059	t complex polypept
40	36	53.7	533	2 T45197	probable serine pr
41	36	53.7	556	2 S13163	t-complex-type mol
42	36	53.7	556	2 S10486	t-complex-type mol
43	36	53.7	556	2 JCI143	t complex polypept
44	36	53.7	556	2 J00866	T-complex protein
45	36	53.7	812	2 A46417	NPI1 protein - yea

ALIGNMENTS

RESULT 1
UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N:Contents: urokinase-type plasminogen activator chain A; urokinase-type plasminogen acti
in form
C:Species: Homo sapiens (man)
C>Date: 17-Dec-1982 #sequence, revision 04-Dec-1986 #text, change 15-Sep-2000
C/Accession: A00931; I52209; J0102; A37561; I8102; S65783; A37563; A37564; A356
R/Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasl, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A>Title: The human urokinase-plasminogen activator gene and its promoter.
A/Reference number: A00931; M01D:85215647; PMID:2987867
A/Accession: A00931
A/Molecule type: DNA
A/Residues: 1-431 <RIC>
A/Cross-references: GB:X02419; NID:G37601; PIDN:CAA26268.1; PID:G1834524
A/Note: the authors translated the codon ATG for residue 214 as Ile
R/Magamine, Y.; Pearson, D.; Gratian, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A>Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine
A/Reference number: I52209; M01D:86050639; PMID:3333505
A/Accession: I52209
A/Status: preliminary; translated from GB/EMBL/DBD
A/Molecule type: DNA
A/Residues: 145-161 <NAG1>
A/Cross-references: GB:X03027; NID:G340174; PIDN:AAA61257.1; PID:G340175
R/Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama, T
Gene 36, 183-188, 1985
A>Title: Molecular cloning of cDNA coding for human preprourokinase.
A/Reference number: J0102; M01D:86056954; PMID:2415429
A/Accession: J0102
A/Molecule type: mRNA
A/Residues: 1-213, 'I', 215-431 <NAG2>
A/Cross-references: GB:X03226; NID:G340155; PIDN:AA97138.1; PID:G340158; GB:D00244; NID
R/Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasl, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A>Title: Identification and primary sequence of an unspliced human urokinase poly(A) + RN
A/Reference number: A37561; M01D:84272706; PMID:6589620
A/Accession: A37561
A/Molecule type: mRNA
A/Residues: 66-431 <VER>
A/Cross-references: GB:D00244; NID:G220138
R/Jacobs, P.; Cravador, A.; Lortiau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen,
DNA 4, 133-146, 1985
A>Title: Molecular cloning, sequencing, and expression in Escherichia coli of human prepr
A/Reference number: I38102; M01D:85203359; PMID:3888571
A/Accession: I38102
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A/Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA26535.1; PID:G35298

R.Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, R. *Biochim. Biophys. Acta* 1293, 83-89, 1996
 A:Title: Characterization of single chain urokinase-type plasminogen activator with a no
 A:Reference number: S65783; MUID:96186279; PMID:8652631
 A:Accession: S65783
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 21-140, 'V', 142-213, 'T', 215-431 <YOS>
 A:Cross-references: EMBL:D1143; NID:9311467; PIDN:BA01919.1; PID:g1199928
 R:Guzlier, W.A.; Steffens, G.U.; Oetting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
 A:Title: The primary structure of high molecular mass urokinase from human urine.
 A:Reference number: A37562; MUID:83055084; PMID:6754569
 A:Accession: A37562
 A:Molecule type: protein
 A:Residues: 21-177 <GUN>
 R:Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Jergler, W.; Studer, R.O.
 Eur. J. Biochem. 125, 251-257, 1982
 A:Title: Human low-molecular-weight urinary urokinase. Partial characterization and preli
 A:Reference number: A37563; MUID:83003608; PMID:6749491
 A:Accession: A37563
 A:Molecule type: protein
 A:Residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>
 R:Steffens, G.U.; Gunzler, W.A.; Oetting, F.; Frankus, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
 A:Title: The complete amino acid sequence of low molecular mass urokinase from human uri
 A:Reference number: A37564; MUID:83055099; PMID:6754572
 A:Accession: A37564
 A:Molecule type: protein
 A:Residues: 158-410 <STE>
 R:Kertzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990
 A:Title: Carboxylate composition and presence of a fucose-protein linkage in recombinat
 A:Reference number: A35689; MUID:90365737; PMID:2393398
 A:Accession: A35689
 A:Molecule type: protein
 A:Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
 A:Note: Identification of a fucose and attempt to determine its attachment site
 R:Rabhand, S.A.; Desjardins, J.; Bell, A.W.; Bannville, D.; Mazur, A.; Henkin, J.; Goltz
 Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
 A:Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
 A:Reference number: A36697; MUID:91097529; PMID:2125213
 A:Accession: A36697
 A:Molecule type: protein
 A:Residues: 21-34 <RAB>
 R:Li, X.; Bokman, A.W.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
 Submitted to the Brookhaven Protein Data Bank, July 1993
 A:Reference number: A51255; PDB:1KDU
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Li, X.; Smith, R.A.G.; Dobson, C.M.
 Biochemistry 31, 9562-9571, 1992
 A:Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain f
 A:Reference number: A44375; MUID:93003110; PMID:1327118
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
 R:Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazur, A.P.; Olejniczak,
 submitted to the Brookhaven Protein Data Bank, January 1994
 A:Reference number: A66822; PDB:1URK
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
 submitted to the Brookhaven Protein Data Bank, July 1995
 A:Reference number: A66058; PDB:1LMM
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175, 179-426
 C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
 C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
 C:Genetics:
 A:Gene: GDB:PLAU
 A:Cross-references: GDB:119497; OMIM:191840
 A:Map position: 10q24-10q24
 A:Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
 C:Function:
 A:Description: proteolytically activates plasminogen
 A:Pathway: fibrinolysis
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteinat
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:1-431/Product: urokinase-type plasminogen activator, single chain form #status predict
 F:1-177/Product: urokinase-type plasminogen activator chain A #status experimental <MPA>
 F:31-62/Domain: EGF homology <EGF>
 F:70-151/Domain: kringle homology <KR>
 F:156-177/Product: urokinase-type plasminogen activator chain B #status experimental <MP
 F:179-431/Product: urokinase-type plasminogen activator chain B #status experimental <MP
 F:179-419/Domain: trypsin homology <TRY>
 F:31-39,33-51,53-62,70-151,91-133,132-146,168-299,209-225,217-288,313-382,345-361,372-400
 F:38/Binding site: carboxylate (Thr) (covalent) #status predicted
 F:178-179/Cleavage site: Lys-116 (plasmin) #status experimental
 F:224,275,376/Active site: His, Asp, Ser #status experimental
 F:322/Binding site: carboxylate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 67; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0.00019;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ASTDTMGRPLP 12
 |||||
 Db 82 ASTDTMGRPLP 93

RESULT 2

UKBAY

u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999

C:Accession: S14687; S08651.
 R:Au, Y.P.T.; Wang, T.W.; Clowes, A.W.

Nucleic Acids Res. 18, 3411, 1990

A:Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen

A:Reference number: S14687; MUID:90287734; PMID:2113276

A:Accession: S14687

A:Molecule type: mRNA

A:Residues: 1-433 <AU>

A:Cross-references: EMBL:X51935; NID:938130; PIDN:CA36200.1; PID:938131

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-176/Product: plasminogen activator chain A #status predicted <ACH>

F:30-61/Domain: EGF homology <EGF>

F:69-150/Domain: kringle homology <KR>

F:178-433/Product: plasminogen activator chain B #status predicted <BC>

F:178-421/Domain: trypsin homology <TRY>

F:167-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted

F:223,274,378/Active site: His, Asp, Ser #status predicted

F:324/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 77.6%; Score 52; DB 1; Length 433;
 Best Local Similarity 90.9%; Pred. No. 0.1;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ASTDTMGRPLP 11
 |||||
 Db 81 ASTDTMGRSCL 91

RESULT 3

S18932

u-plasminogen activator (EC 3.4.21.73) precursor - rat

N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator

C:Species: Rattus norvegicus (Norway rat)

C:Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999

C:Accession: S24604; I60186; I53472; S18932

R:Rabhand, S.A.

submitted to the EMBL Data Library, April 1992

A:Reference number: S24604

A:Accession: S24604

A:Molecule type: mRNA

A:Residues: 1-15, 'H', 17-23, 'G', 25-33, 'N', 33-432 <RAB>

A:Cross-references: EMBL:X65651; NID:957456; PIDN:CAA4601.1; PID:957457

A:Experimental source: tissue kidney
 R:Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
 Cancer Res. 52, 2489-2496, 1992
 A>Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
 A:Reference number: 160186; MUID:92233409; PMID:1568219
 A:Accession: 160186
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-432 <RES>
 A:Cross-references: EMBL:X63434; NID:957465; PIDN:CAA45028.1; PID:957466
 A:Experimental source: strain Fischer 344; tissue mammary
 R:Ragno, F.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
 FEBS Lett. 306, 193-198, 1992
 A>Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
 A:Reference number: 153472; MUID:92339549; PMID:1321734
 A:Accession: 153472
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 31-62 <RES2>
 A:Cross-references: EMBL:X66907; NID:9396200; PIDN:CAA47356.1; PID:9398279
 C:Genetics: uRA
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:31-62/Domain: EGF homology <EGF>
 F:70-151/Domain: kringle homology <KRG>
 F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:179-420/Domain: trypsin homology <TRY>
 F:168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted
 F:225,276,377/Active site: His, Asp, Ser #status predicted

Query Match 76.1%; Score 51; DB 1; Length 432;
 Best Local Similarity 81.8%; Pred. No. 0.16;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCL 11
 Db 82 ANNTDGRPCL 92

RESULT 4
 UKMS
 u-plasminogen activator (EC 3.4.21.73) precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
 C:Accession: A29420; A24615
 R:Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
 Biochemistry 26, 8270-8279, 1987
 A>Title: The murine urokinase-type plasminogen activator gene.
 A:Reference number: A29420; MUID:86163489; PMID:2831940
 A:Accession: A29420
 A:Molecule type: DNA
 A:Residues: 1-433 <DEG>
 A:Cross-references: GB:M17922; NID:9202296; PIDN:AAA40539.1; PID:9202297
 R:Belin, D.; Vassalli, J.D.; Combepine, C.; Godreau, F.; Nagamine, Y.; Reich, E.; Kocher,
 Eur. J. Biochem. 148, 225-232, 1995
 A>Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
 A:Reference number: A24615; MUID:85179474; PMID:2985383
 A:Accession: A24615
 A:Molecule type: mRNA
 A:Residues: 1-433 <BEL>
 A:Cross-references: GB:X02389; NID:955127; PIDN:CAA2631.1; PID:955128
 C:Genetics: uRA
 A:Initons: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:32-65/Domain: EGF homology <EGF>
 F:71-152/Domain: kringle homology <KRG>
 F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

F:180-421/Domain: trypsin homology <TRY>
 F:159-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 76.1%; Score 51; DB 1; Length 433;
 Best Local Similarity 81.8%; Pred. No. 0.16;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCL 11
 Db 83 ANNTDGRPCL 93

RESULT 5
 UKPG
 u-plasminogen activator (EC 3.4.21.73) precursor - pig
 N:Alternate names: uPA
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C:Accession: A00932
 R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A>Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A:Reference number: A00932; MUID:85087954; PMID:6096832
 A:Accession: A00932
 A:Molecule type: DNA
 A:Residues: 1-240, 'H', 242-442 <NAG1>
 A:Experimental source: kidney cell line LLC-PK1
 R:Nagamine, Y.
 submitted to the Protein Sequence Database, December 1986
 A:Reference number: A37566
 A:Contents: annotation; correction to residue 241
 C:Genetics: uRA
 A:Initons: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:33-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KRG>
 F:150-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:190-430/Domain: trypsin homology <TRY>
 F:152/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:179-310,220-236,228-299,334-393,356-372,383-411/Disulfide bonds: #status predicted
 F:235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 76.1%; Score 51; DB 1; Length 442;
 Best Local Similarity 75.0%; Pred. No. 0.16;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCL 12
 Db 84 ANNTDGRPCL 95

RESULT 6
 S20463
 siderophore biosynthesis regulatory protein sfp(0) - Bacillus subtilis
 N:Alternate names: surfactin producing protein
 C:Species: Bacillus subtilis
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Jun-2000
 C:Accession: S20463; A36931; F69705; S26594
 R:Nakano, M.M.; Corbell, N.; Besson, J.; Zuber, P.
 Mol. Gen. Genet. 232, 313-321, 1992
 A>Title: Isolation and characterization of sfp: a gene that functions in the production
 A:Reference number: S20463; MUID:92212296; PMID:1557038
 A:Accession: S20463
 A:Molecule type: DNA
 A:Residues: 1-224 <NAK>
 A:Cross-references: EMBL:X63158; NID:940138; PIDN:CAA44858.1; PID:940139
 A:Experimental source: strain OKB105
 R:Grossman, T.H.; Tuckman, M.; Ellestead, S.; Osburne, M.S.
 J. Bacteriol. 175, 6203-6211, 1993

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Nov-1999
 C/Accession: S42989
 R/Strut: D.I.; White, R.A.H.
 A/Description: Characterisation of T48, a target of homeotic gene regulation in *Drosophila*
 A/Reference number: S42989
 A/Accession: S42989
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-419 <STR>
 A/Cross-references: EMBL:X78113; NID:G459593; PID:G459594
 A/Gene: FLYBase:T48
 A/Cross-references: FLYBase:FBgn0004359
 C/Superfamily: fruit fly T48 protein

Query Match 58.2%; Score 39; DB 2; Length 419;
 Best Local Similarity 58.3%; Pred. No. 24;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ASTDTMGRPCLP 12
 Db 135 AHTEFMGRVCP 146

RESULT 11
 JN0560
 U-plasminogen activator (EC 3.4.21.73) precursor - bovine
 N/Alternate names: uPA
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C/Accession: JN0560
 R/Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schlemming, W.D.
 Gene 125, 177-183, 1993
 A/Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induction
 A/Reference number: JN0560; MUID:93216119; PMID:8385052
 A/Accession: JN0560
 A/Molecule type: mRNA
 A/Residues: 1-433 <KRA>
 A/Cross-references: GB:103546; NID:G163800; PID:AAA51419.1; PID:G163801
 C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine protease
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-119/Product: plasminogen activator chain A #status predicted <MA1>
 F:21-119/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:33-66/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KR>
 F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
 F:181-421/Domain: trypsin homology <TRY>
 F:170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 58.2%; Score 39; DB 1; Length 433;
 Best Local Similarity 63.6%; Pred. No. 25;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ASTDTMGRPCLP 11
 Db 84 ANRDLGRPCL 94

RESULT 12
 AF1927
 hypothetical protein al10969 [imported] - Nostoc sp. (strain PCC 7120)
 C/Species: Nostoc sp. PCC 7120
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C/Accession: AF1927
 R/Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tsubata, S.
 DNA Res. 8, 205-213, 2001
 A/Title: Complete Genome Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*
 A/Reference number: AB1807; MUID:21595285; PMID:11759840

A/Accession: AF1927
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-543 <KUR>
 A/Cross-references: GB:PA000019; PIDN:BA872926.1; PID:G17130315; GSPDB:GN00179
 A/Experimental source: strain PCC 7120
 C/Genetics: A/Gene: al10969

Query Match 58.2%; Score 39; DB 2; Length 543;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ASTDTMGRPCLP 12
 Db 230 ANTDTCQPPFP 241

RESULT 13
 A35827
 thrombin (EC 3.4.21.5) precursor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 03-May-2002
 C/Accession: A35827; A42696; S12081
 R/Degen, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgerald, J.J.; Pal, J.A.
 DNA Cell Biol. 9, 487-498, 1990
 A/Title: Characterization of the cDNA coding for mouse prothrombin and localization of th
 A/Reference number: A35827; MUID:91025551; PMID:2222810
 A/Accession: A35827
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-618 <DEG>
 A/Cross-references: GB:X52308; NID:G53813; PIDN:CAA36548.1; PID:G53814
 A/Experimental source: strain C57BL/6
 A/Note: the data were obtained from females resulting from the cross of M. domesticus and
 R/Banfield, D.K.; MacGillivray, R.T.A.
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
 A/Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequ
 A/Reference number: A42696; MUID:92212913; PMID:1557383
 A/Accession: A42696
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 384-618, 'E' <BAN>
 A/Cross-references: GB:M81394
 C/Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
 C/Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; hydr
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-43/Domain: propeptide #status predicted <PRO>
 F:28-88/Domain: Gla domain homology <GLA>
 F:44-618/Product: prothrombin B #status predicted <MAT>
 F:109-187/Domain: kringle homology <KR1>
 F:215-293/Domain: kringle homology <KR2>
 F:361-610/Domain: trypsin homology <TRY>
 F:50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status i
 F:61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-547,56
 F:403,459,565/Active site: His, Asp, Ser #status predicted

Query Match 58.2%; Score 39; DB 2; Length 618;
 Best Local Similarity 75.0%; Pred. No. 35;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 TMGRPCLP 12
 Db 231 TLGSPCLP 238

RESULT 14
 T49092
 5B-like protein - Arabidopsis thaliana
 N/Alternate names: protein P4P15.240
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
 C/Accession: T49092

R;Alcaraz, J.P.; Clabault, G.; Cottet, A.; Mache, R.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25015
A;Accession: T49092
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-125 <ALC>
A;Cross-references: EMBL:AL049711, GSPDB:GN00061, ATSP:FAF15.240
A;Experimental source: cultivar Columbia; BAC clone FAF15
C;Genetics:
A;Gene: ATSP:FAF15.240
A;Map position: 3

Query Match 56.7%; Score 38; DB 2; Length 125;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 5 TMGRPCL 11
|:|||||
Db 84 TLGRPCL 90

RESULT 15

H72579
hypothetical protein APE1920 - Aeropyrum pernix (strain KI)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: H72579
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999.
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: H72579
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-167 <KAW>
A;Cross-references: DDBJ:AP000062; NID:G5105244; PIDN:BAA80925.1; PID:dl044711; PID:G510
A;Experimental source: strain KI
C;Genetics:
A;Gene: APE1920

Query Match 56.7%; Score 38; DB 2; Length 167;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 TMGRPCLP 12
|:|||||
Db 64 TFGRPCLP 71

Search completed: March 18, 2004, 11:20:18
Job time : 2.23414 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:54:45 ; Search time 0.73523 Seconds
(without alignments)
849.859 Million cell updates/sec

Title: US-10-076-421-4
Perfect score: 67
Sequence: 1 ASTDTNGRPCLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	431	1 UROK_HUMAN	P00749 homo sapien
2	52	77.6	433	1 UROK_PAPCY	P16227 papio cynoc
3	51	76.1	432	1 UROK_RAT	P29598 rattus norv
4	51	76.1	433	1 UROK_MOUSE	P06869 mus musculu
5	51	76.1	442	1 UROK_PIG	P04185 sus scrofa
6	40	59.7	224	1 SFP_BACSU	P39135 bacillus su
7	40	59.7	500	1 CPDQ_CAVPO	O64403 cavia porce
8	39	58.2	433	1 UROK_BOVIN	O05589 bos taurus
9	39	58.2	602	1 YHS4_CABEL	O18363 caenorhabdi
10	38	58.2	618	1 THRB_MOUSE	P19221 mus musculu
11	38	56.7	585	1 PTRR_PIG	P50133 sus scrofa
12	38	56.7	591	1 PTRR_MOUSE	P41593 mus musculu
13	38	56.7	591	1 PTRR_RAT	P25961 rattus norv
14	38	56.7	593	1 PTRR_HUMAN	O03431 homo sapien
15	37	55.2	205	1 PSB3_TRYBB	O9nd81 trypanosoma
16	37	55.2	376	1 VIBX_AGRTR	P09789 agrobacteri
17	37	55.2	377	1 VIBX_AGRTR9	P05359 agrobacteri
18	37	55.2	810	1 HTR4_HALNI	O9hpd4 halobacteri
19	37	55.2	810	1 HTR4_HALSA	O48317 halobacteri
20	37	55.2	1299	1 VP75_HSVSA	P11282 herpesviru
21	37	55.2	1693	1 SAS_DROME	O41464 drosophila
22	36	53.7	230	1 EIA_ADECC	O65941 canine aden
23	36	53.7	247	1 YE95_SCHPO	O13767 schizosacch
24	36	53.7	353	1 NIFA_RHILT	P24436 rhizobium
25	36	53.7	374	1 ZIP4_ARATH	O04089 arabidopsis
26	36	53.7	395	1 EFTU_XYLEA	O9p9g9 xylella fas
27	36	53.7	457	1 TMS5_HUMAN	O9h333 homo sapien
28	36	53.7	457	1 TMS5_HUMAN	O9h333 homo sapien
29	36	53.7	556	1 TCP2_MOUSE	P18293 mus musculu
30	36	53.7	556	1 TCPA_CRIGR	P17967 homo sapien
31	36	53.7	556	1 TCPA_HUMAN	P28480 rattus norv
32	36	53.7	556	1 TCPA_RAT	P28480 rattus norv
33	36	53.7	559	1 TCPA_PALPA	O9w790 paleosuchus

34	36	53.7	812	1 IF38_YEAST	P32497 saccharomyc
35	36	53.7	864	1 AGLU_MUCJA	O92442 mucor javan
36	36	53.7	954	1 YB79_YEAST	P38138 saccharomyc
37	36	53.7	3164	1 TEGU_HSV11	P10220 herpes simp
38	35	52.2	126	1 MN72_EPRST	P28123 eptaretus
39	35	52.2	224	1 FFP_BACSU	O96447 bacillus su
40	35	52.2	477	1 NIFB_RHOCA	P19055 rhodobacter
41	35	52.2	530	1 CP14_CHICK	P39760 gallus galli
42	35	52.2	630	1 KCD2_HUMAN	O9azv8 homo sapien
43	35	52.2	690	1 AFT1_YEAST	P22149 saccharomyc
44	35	52.2	696	1 SYM_XANCP	O8pay7 xanthomonas
45	35	52.2	726	1 CNG2_ARATH	O65718 arabidopsis

ALIGNMENTS

RESULT 1
UROK_HUMAN STANDARD; PRT; 431 AA.
AC P00749; Q15844; Q1618; Q699W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (BC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=85215647; PubMed=2987867;
RX "Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RT Steffens G.J., Heyneker H.L.,
RT "Cloning and expression of the gene for pro-urokinase in Escherichia
colli.";
RL Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=6056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
RT "Molecular cloning of cDNA coding for human prepro-urokinase.";
RL Gene 36:183-188(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Cravador A., Lorlau R., Brockly F., Colau B., Chuchana P.,
RT "Molecular cloning, sequencing, and expression in Escherichia coli of
human prepro-urokinase cDNA.";
RL DNA 4:139-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi O.,
RT Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Stranberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT Klausner R.D., Collins F.S., Wagner L., Stenmen C.M., Shuler G.D.,
RA Altschuler S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heileh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Pringle C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RX SEQUENCE OF 66-431 FROM N.A.
 RA MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blaesi F.,
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [8]
 RX SEQUENCE OF 21-177.
 RA MEDLINE=83055084; PubMed=6754569;
 RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.,
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [9]
 RX SEQUENCE OF 156-176 AND 179-224.
 RA MEDLINE=83003608; PubMed=6749921;
 RA Schaller U., Nick H., Rickli E.E., Gillissen D., Lergier W.,
 RA Struder R.O.,
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RL Eur. J. Biochem. 125:251-257(1982).
 RN [10]
 RX SEQUENCE OF 158-410.
 RA MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.,
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN [11]
 RX X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RA MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.,
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RL Structure 3:681-691(1995).
 RN [12]
 RX X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RA MEDLINE=20266227; PubMed=10805774;
 RA Spehl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdolen V., Huber R., Moroder L.,
 RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13]
 RX STRUCTURE BY NMR.
 RA MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Boguski M.J., Bamberger M., Smith R.A.G., Dobson C.M.,
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14]
 RX STRUCTURE BY NMR OF 67-155.
 RA MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.,

RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15]
 RX STRUCTURE BY NMR OF 67-155.
 RA MEDLINE=9419701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.,
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RX VARIANT LEU-141.
 RA MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
 RA Sawasaki Y., Hanada K.,
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle
 RT structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RX VARIANT LEU-141.
 RA MEDLINE=97218551; PubMed=9065988;
 RA Comme B., Berczy M., Belin D.,
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RX ERRATUM.
 RA Comme B., Berczy M., Belin D.,
 RL Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RX VARIANT LEU-141.
 RA MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Heil W.,
 RA Creutzburg S., Graef H., Magdolen V.,
 RT "Mutational analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -1- FUNCTION: Potent plasminogen activator and is clinically used for
 CC therapy of thrombolytic disorders.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1 chain.
 CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
 CC in pulmonary embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL, X02419; CAA26268.1; -;
 DR EMBL, M15476; AAA61253.1; -;
 DR EMBL, D00244; BAA00175.1; -;
 DR EMBL, D11143; BAA01919.1; -;
 DR EMBL, X02760; CAA26535.1; -;
 DR EMBL, AF377330; AAK53822.1; -;
 DR EMBL, BC013575; AAH13575.1; -;
 DR EMBL, K03226; AAC97138.1; -;
 DR EMBL, K02286; AAA61252.1; -;
 DR EMBL, A21571; CAA01559.1; -;
 DR EMBL, A18397; CAA01390.1; -;
 DR PIR, A00931; UKHD.
 DR PDB, 1KDU; 3I-OCT-93.

Query Match 100.0%; Score 67; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 6e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASTDTMGRPCLP 12
 DB 82 ASTDTMGRPCLP 93

RESULT 2
 UROK_PAPCY STANDARD; PRT; 433 AA.

AC P16227;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Papio cynocephalus (Yellow baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Papio.
 OX NCBI_TaxID=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thoracic aorta;
 RX MEDLINE=90287734; PubMed=2113276;
 RA Au Y.P.T., Wang T.W., Clowes A.W.;
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
 RT plasminogen activator."
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1-chain (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

EMBL: X51935; CAA36200.1; -.
 DR PIR; S14687; UKBAY.
 DR HSSP; P00749; ILMW.
 DR MEROPS; S01.231; -.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR008293; Pept_S1A_uPA.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PIRSF; PIRSF001144; UK_plasm_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_Spec; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.

DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KM plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL
 FT CHAIN 1
 FT CHAIN 20
 FT CHAIN 21 433
 FT CHAIN 21 176
 FT CHAIN 155 176
 FT CHAIN 178 433
 FT CHAIN 26 62
 FT DOMAIN 69 150
 FT DOMAIN 151 177
 FT DOMAIN 178 433
 FT DISULFID 30 38
 FT DISULFID 32 50
 FT DISULFID 52 61
 FT DISULFID 167 298
 FT DISULFID 208 224
 FT DISULFID 216 287
 FT DISULFID 315 384
 FT DISULFID 347 363
 FT DISULFID 374 402
 FT ACT_SITE 223 223
 FT ACT_SITE 274 274
 FT ACT_SITE 378 378
 FT CARBOHYD 324 324
 SQ SEQUENCE 433 AA; 48595 MW; 816D2DPEDDC8792 CRC64; (BY SIMILARITY).

Query Match 77.6%; Score 52; DB 1; Length 433;
 Best Local Similarity 90.9%; Pred. No. 0.038;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ASTDTMGRPCLP 11
 DB 81 ASTDTMGRPCLP 91

RESULT 3
 UROK_RAT STANDARD; PRT; 432 AA.

AC P29598;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer 344;
 RX MEDLINE=92233409; PubMed=1568219;
 RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
 RA Kefford R.F.;
 RT "Transcriptional and posttranscriptional activation of urokinase
 RT plasminogen activator gene expression in metastatic tumor cells."
 RT Cancer Res. 52:2489-2496(1992).
 RL [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Rabbani S.A.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a

CC long chain A. Cleavage occurs after residue 156 in the low
 CC molecular mass form to yield a short A1 chain (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X63434; CAA45028.1; -;
 DR EMBL; X65651; CAA46601.1; -;
 DR PIR; S24604; S18932.
 DR HSSP; P00749; 1KDU.
 DR MEROPS; S01.231; -;
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR008293; Pept_S1A_uPA.
 DR InterPro; IPR01254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_Spec; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 20 177 CHAIN A (BY SIMILARITY).
 FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
 FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
 FT DOMAIN 27 63 EGF-LIKE.
 FT DOMAIN 70 151 KRINGLE.
 FT DOMAIN 152 178 CONNECTING PEPTIDE.
 FT DOMAIN 179 432 SERINE PROTEASE.
 FT DISULFID 31 39 BY SIMILARITY.
 FT DISULFID 33 51 BY SIMILARITY.
 FT DISULFID 53 62 BY SIMILARITY.
 FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 210 226 BY SIMILARITY.
 FT DISULFID 218 289 BY SIMILARITY.
 FT DISULFID 314 383 BY SIMILARITY.
 FT DISULFID 346 362 BY SIMILARITY.
 FT DISULFID 373 401 BY SIMILARITY.
 FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
 FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.
 FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
 FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
 FT CONFLICT 16 16 N -> H (IN REF. 2).
 FT CONFLICT 24 24 E -> G (IN REF. 2).
 FT CONFLICT 332 332 D -> N (IN REF. 2).
 SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 ASTDTMGRPCL 11
 Db 82 ANTDTKGRPCL 92
 RESULT 4
 ID UROK_MOUSE STANDARD; PRT; 433 AA.
 AC P06863;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (BC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85179474; PubMed=2985383;
 RA Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,
 RA Reich E., Koerber H.P., Duvoisin R.M.;
 RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
 RT mouse urokinase-type plasminogen activator.";
 RL Eur. J. Biochem. 148:225-232(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88163489; PubMed=2831940;
 RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
 RT "The murine urokinase-type plasminogen activator gene.";
 RL Biochemistry 26:8270-8279(1987).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 156 in the low
 CC molecular mass form to yield a short A1 chain (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02389; CAA26231.1; -;
 DR EMBL; M17922; AAA40539.1; -;
 DR PIR; A29420; UROKS.
 DR HSSP; P00749; 1KDU.
 DR MEROPS; S01.231; -;
 DR MGD; MGI.97611; Plau.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR01254; Pept_S1A_uPA.
 DR InterPro; IPR001314; Peptidase_S1.
 DR InterPro; IPR001254; Peptidase_S1A.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PIRSF; PIRSF001144; Urk_plasm_act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.

RESULT 6
SFP_BACSU STANDARD, PRT, 224 AA.
ID_SFP_BACSU PRT, 224 AA.
AC P39135;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 4'-phosphopantetheinyl transferase sfp (EC 2.7.8.-) (Surfactin synthetase-activating enzyme).
GN Sfp OR LPA-8 OR BSU03570.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxId=1423;
[1]
[2]
[3]
[4]
[5]
[6]
[7]
[8]
[9]
[10]
[11]
[12]
[13]
[14]
[15]
[16]
[17]
[18]
[19]
[20]
[21]
[22]
[23]
[24]
[25]
[26]
[27]
[28]
[29]
[30]
[31]
[32]
[33]
[34]
[35]
[36]
[37]
[38]
[39]
[40]
[41]
[42]
[43]
[44]
[45]
[46]
[47]
[48]
[49]
[50]
[51]
[52]
[53]
[54]
[55]
[56]
[57]
[58]
[59]
[60]
[61]
[62]
[63]
[64]
[65]
[66]
[67]
[68]
[69]
[70]
[71]
[72]
[73]
[74]
[75]
[76]
[77]
[78]
[79]
[80]
[81]
[82]
[83]
[84]
[85]
[86]
[87]
[88]
[89]
[90]
[91]
[92]
[93]
[94]
[95]
[96]
[97]
[98]
[99]
[100]
[101]
[102]
[103]
[104]
[105]
[106]
[107]
[108]
[109]
[110]
[111]
[112]
[113]
[114]
[115]
[116]
[117]
[118]
[119]
[120]
[121]
[122]
[123]
[124]
[125]
[126]
[127]
[128]
[129]
[130]
[131]
[132]
[133]
[134]
[135]
[136]
[137]
[138]
[139]
[140]
[141]
[142]
[143]
[144]
[145]
[146]
[147]
[148]
[149]
[150]
[151]
[152]
[153]
[154]
[155]
[156]
[157]
[158]
[159]
[160]
[161]
[162]
[163]
[164]
[165]
[166]
[167]
[168]
[169]
[170]
[171]
[172]
[173]
[174]
[175]
[176]
[177]
[178]
[179]
[180]
[181]
[182]
[183]
[184]
[185]
[186]
[187]
[188]
[189]
[190]
[191]
[192]
[193]
[194]
[195]
[196]
[197]
[198]
[199]
[200]
[201]
[202]
[203]
[204]
[205]
[206]
[207]
[208]
[209]
[210]
[211]
[212]
[213]
[214]
[215]
[216]
[217]
[218]
[219]
[220]
[221]
[222]
[223]
[224]
[225]
[226]
[227]
[228]
[229]
[230]
[231]
[232]
[233]
[234]
[235]
[236]
[237]
[238]
[239]
[240]
[241]
[242]
[243]
[244]
[245]
[246]
[247]
[248]
[249]
[250]
[251]
[252]
[253]
[254]
[255]
[256]
[257]
[258]
[259]
[260]
[261]
[262]
[263]
[264]
[265]
[266]
[267]
[268]
[269]
[270]
[271]
[272]
[273]
[274]
[275]
[276]
[277]
[278]
[279]
[280]
[281]
[282]
[283]
[284]
[285]
[286]
[287]
[288]
[289]
[290]
[291]
[292]
[293]
[294]
[295]
[296]
[297]
[298]
[299]
[300]
[301]
[302]
[303]
[304]
[305]
[306]
[307]
[308]
[309]
[310]
[311]
[312]
[313]
[314]
[315]
[316]
[317]
[318]
[319]
[320]
[321]
[322]
[323]
[324]
[325]
[326]
[327]
[328]
[329]
[330]
[331]
[332]
[333]
[334]
[335]
[336]
[337]
[338]
[339]
[340]
[341]
[342]
[343]
[344]
[345]
[346]
[347]
[348]
[349]
[350]
[351]
[352]
[353]
[354]
[355]
[356]
[357]
[358]
[359]
[360]
[361]
[362]
[363]
[364]
[365]
[366]
[367]
[368]
[369]
[370]
[371]
[372]
[373]
[374]
[375]
[376]
[377]
[378]
[379]
[380]
[381]
[382]
[383]
[384]
[385]
[386]
[387]
[388]
[389]
[390]
[391]
[392]
[393]
[394]
[395]
[396]
[397]
[398]
[399]
[400]
[401]
[402]
[403]
[404]
[405]
[406]
[407]
[408]
[409]
[410]
[411]
[412]
[413]
[414]
[415]
[416]
[417]
[418]
[419]
[420]
[421]
[422]
[423]
[424]
[425]
[426]
[427]
[428]
[429]
[430]
[431]
[432]
[433]
[434]
[435]
[436]
[437]
[438]
[439]
[440]
[441]
[442]
[443]
[444]
[445]
[446]
[447]
[448]
[449]
[450]
[451]
[452]
[453]
[454]
[455]
[456]
[457]
[458]
[459]
[460]
[461]
[462]
[463]
[464]
[465]
[466]
[467]
[468]
[469]
[470]
[471]
[472]
[473]
[474]
[475]
[476]
[477]
[478]
[479]
[480]
[481]
[482]
[483]
[484]
[485]
[486]
[487]
[488]
[489]
[490]
[491]
[492]
[493]
[494]
[495]
[496]
[497]
[498]
[499]
[500]
[501]
[502]
[503]
[504]
[505]
[506]
[507]
[508]
[509]
[510]
[511]
[512]
[513]
[514]
[515]
[516]
[517]
[518]
[519]
[520]
[521]
[522]
[523]
[524]
[525]
[526]
[527]
[528]
[529]
[530]
[531]
[532]
[533]
[534]
[535]
[536]
[537]
[538]
[539]
[540]
[541]
[542]
[543]
[544]
[545]
[546]
[547]
[548]
[549]
[550]
[551]
[552]
[553]
[554]
[555]
[556]
[557]
[558]
[559]
[560]
[561]
[562]
[563]
[564]
[565]
[566]
[567]
[568]
[569]
[570]
[571]
[572]
[573]
[574]
[575]
[576]
[577]
[578]
[579]
[580]
[581]
[582]
[583]
[584]
[585]
[586]
[587]
[588]
[589]
[590]
[591]
[592]
[593]
[594]
[595]
[596]
[597]
[598]
[599]
[600]
[601]
[602]
[603]
[604]
[605]
[606]
[607]
[608]
[609]
[610]
[611]
[612]
[613]
[614]
[615]
[616]
[617]
[618]
[619]
[620]
[621]
[622]
[623]
[624]
[625]
[626]
[627]
[628]
[629]
[630]
[631]
[632]
[633]
[634]
[635]
[636]
[637]
[638]
[639]
[640]
[641]
[642]
[643]
[644]
[645]
[646]
[647]
[648]
[649]
[650]
[651]
[652]
[653]
[654]
[655]
[656]
[657]
[658]
[659]
[660]
[661]
[662]
[663]
[664]
[665]
[666]
[667]
[668]
[669]
[670]
[671]
[672]
[673]
[674]
[675]
[676]
[677]
[678]
[679]
[680]
[681]
[682]
[683]
[684]
[685]
[686]
[687]
[688]
[689]
[690]
[691]
[692]
[693]
[694]
[695]
[696]
[697]
[698]
[699]
[700]
[701]
[702]
[703]
[704]
[705]
[706]
[707]
[708]
[709]
[710]
[711]
[712]
[713]
[714]
[715]
[716]
[717]
[718]
[719]
[720]
[721]
[722]
[723]
[724]
[725]
[726]
[727]
[728]
[729]
[730]
[731]
[732]
[733]
[734]
[735]
[736]
[737]
[738]
[739]
[740]
[741]
[742]
[743]
[744]
[745]
[746]
[747]
[748]
[749]
[750]
[751]
[752]
[753]
[754]
[755]
[756]
[757]
[758]
[759]
[760]
[761]
[762]
[763]
[764]
[765]
[766]
[767]
[768]
[769]
[770]
[771]
[772]
[773]
[774]
[775]
[776]
[777]
[778]
[779]
[780]
[781]
[782]
[783]
[784]
[785]
[786]
[787]
[788]
[789]
[790]
[791]
[792]
[793]
[794]
[795]
[796]
[797]
[798]
[799]
[800]
[801]
[802]
[803]
[804]
[805]
[806]
[807]
[808]
[809]
[810]
[811]
[812]
[813]
[814]
[815]
[816]
[817]
[818]
[819]
[820]
[821]
[822]
[823]
[824]
[825]
[826]
[827]
[828]
[829]
[830]
[831]
[832]
[833]
[834]
[835]
[836]
[837]
[838]
[839]
[840]
[841]
[842]
[843]
[844]
[845]
[846]
[847]
[848]
[849]
[850]
[851]
[852]
[853]
[854]
[855]
[856]
[857]
[858]
[859]
[860]
[861]
[862]
[863]
[864]
[865]
[866]
[867]
[868]
[869]
[870]
[871]
[872]
[873]
[874]
[875]
[876]
[877]
[878]
[879]
[880]
[881]
[882]
[883]
[884]
[885]
[886]
[887]
[888]
[889]
[890]
[891]
[892]
[893]
[894]
[895]
[896]
[897]
[898]
[899]
[900]
[901]
[902]
[903]
[904]
[905]
[906]
[907]
[908]
[909]
[910]
[911]
[912]
[913]
[914]
[915]
[916]
[917]
[918]
[919]
[920]
[921]
[922]
[923]
[924]
[925]
[926]
[927]
[928]
[929]
[930]
[931]
[932]
[933]
[934]
[935]
[936]
[937]
[938]
[939]
[940]
[941]
[942]
[943]
[944]
[945]
[946]
[947]
[948]
[949]
[950]
[951]
[952]
[953]
[954]
[955]
[956]
[957]
[958]
[959]
[960]
[961]
[962]
[963]
[964]
[965]
[966]
[967]
[968]
[969]
[970]
[971]
[972]
[973]
[974]
[975]
[976]
[977]
[978]
[979]
[980]
[981]
[982]
[983]
[984]
[985]
[986]
[987]
[988]
[989]
[990]
[991]
[992]
[993]
[994]
[995]
[996]
[997]
[998]
[999]
[1000]

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Punelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scifone F., Soldo B.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Sordo B.,
RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vanlier F., Vassarotti A.,
RA Viari A., Wambuit R., Wedler E., Weitzenger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*.";
RT Nature 390:249-256 (1997).
[7]
[8]
[9]
[10]
[11]
[12]
[13]
[14]
[15]
[16]
[17]
[18]
[19]
[20]
[21]
[22]
[23]
[24]
[25]
[26]
[27]
[28]
[29]
[30]
[31]
[32]
[33]
[34]
[35]
[36]
[37]
[38]
[39]
[40]
[41]
[42]
[43]
[44]
[45]
[46]
[47]
[48]
[49]
[50]
[51]
[52]
[53]
[54]
[55]
[56]
[57]
[58]
[59]
[60]
[61]
[62]
[63]
[64]
[65]
[66]
[67]
[68]
[69]
[70]
[71]
[72]
[73]
[74]
[75]
[76]
[77]
[78]
[79]
[80]
[81]
[82]
[83]
[84]
[85]
[86]
[87]
[88]
[89]
[90]
[91]
[92]
[93]
[94]
[95]
[96]
[97]
[98]
[99]
[100]
[101]
[102]
[103]
[104]
[105]
[106]
[107]
[108]
[109]
[110]
[111]
[112]
[113]
[114]
[115]
[116]
[117]
[118]
[119]
[120]
[121]
[122]
[123]
[124]
[125]
[126]
[127]
[128]
[129]
[130]
[131]
[132]
[133]
[134]
[135]
[136]
[137]
[138]
[139]
[140]
[141]
[142]
[143]
[144]
[145]
[146]
[147]
[148]
[149]
[150]
[151]
[152]
[153]
[154]
[155]
[156]
[157]
[158]
[159]
[160]
[161]
[162]
[163]
[164]
[165]
[166]
[167]
[168]
[169]
[170]
[171]
[172]
[173]
[174]
[175]
[176]
[177]
[178]
[179]
[180]
[181]
[182]
[183]
[184]
[185]
[186]
[187]
[188]
[189]
[190]
[191]
[192]
[193]
[194]
[195]
[196]
[197]
[198]
[199]
[200]
[201]
[202]
[203]
[204]
[205]
[206]
[207]
[208]
[209]
[210]
[211]
[212]
[213]
[214]
[215]
[216]
[217]
[218]
[219]
[220]
[221]
[222]
[223]
[224]
[225]
[226]
[227]
[228]
[229]
[230]
[231]
[232]
[233]
[234]
[235]
[236]
[237]
[238]
[239]
[240]
[241]
[242]
[243]
[244]
[245]
[246]
[247]
[248]
[249]
[250]
[251]
[252]
[253]
[254]
[255]
[256]
[257]
[258]
[259]
[260]
[261]
[262]
[263]
[264]
[265]
[266]
[267]
[268]
[269]
[270]
[271]
[272]
[273]
[274]
[275]
[276]
[277]
[278]
[279]
[280]
[281]
[282]
[283]
[284]
[285]
[286]
[287]
[288]
[289]
[290]
[291]
[292]
[293]
[294]
[295]
[296]
[297]
[298]
[299]
[300]
[301]
[302]
[303]
[304]
[305]
[306]
[307]
[308]
[309]
[310]
[311]
[312]
[313]
[314]
[315]
[316]
[317]
[318]
[319]
[320]
[321]
[322]
[323]
[324]
[325]
[326]
[327]
[328]
[329]
[330]
[331]
[332]
[333]
[334]
[335]
[336]
[337]
[338]
[339]
[340]
[341]
[342]
[343]
[344]
[345]
[346]
[347]
[348]
[349]
[350]
[351]
[352]
[353]
[354]
[355]
[356]
[357]
[358]
[359]
[360]
[361]
[362]
[363]
[364]
[365]
[366]
[367]
[368]
[369]
[370]
[371]
[372]
[373]
[374]
[375]
[376]
[377]
[378]
[379]
[380]
[381]
[382]
[383]
[384]
[385]
[386]
[387]
[388]
[389]
[390]
[391]
[392]
[393]
[394]
[395]
[396]
[397]
[398]
[399]
[400]
[401]
[402]
[403]
[404]
[405]
[406]
[407]
[408]
[409]
[410]
[411]
[412]
[413]
[414]
[415]
[416]
[417]
[418]
[419]
[420]
[421]
[422]
[423]
[424]
[425]
[426]
[427]
[428]
[429]
[430]
[431]
[432]
[433]
[434]
[435]
[436]
[437]
[438]
[439]
[440]
[441]
[442]
[443]
[444]
[445]
[446]
[447]
[448]
[449]
[450]
[451]
[452]
[453]
[454]
[455]
[456]
[457]
[458]
[459]
[460]
[461]
[462]
[463]
[464]
[465]
[466]
[467]
[468]
[469]
[470]
[471]
[472]
[473]
[474]
[475]
[476]
[477]
[478]
[479]
[480]
[481]
[482]
[483]
[484]
[485]
[486]
[487]
[488]
[489]
[490]
[491]
[492]
[493]
[494]
[495]
[496]
[497]
[498]
[499]
[500]
[501]
[502]
[503]
[504]
[505]
[506]
[507]
[508]
[509]
[510]
[511]
[512]
[513]
[514]
[515]
[516]
[517]
[518]
[519]
[520]
[521]
[522]
[523]
[524]
[525]
[526]
[527]
[528]
[529]
[530]
[531]
[532]
[533]
[534]
[535]
[536]
[537]
[538]
[539]
[540]
[541]
[542]
[543]
[544]
[545]
[546]
[547]
[548]
[549]
[550]
[551]
[552]
[553]
[554]
[555]
[556]
[557]
[558]
[559]
[560]
[561]
[562]
[563]
[564]
[565]
[566]
[567]
[568]
[569]
[570]
[571]
[572]
[573]
[574]
[575]
[576]
[577]
[578]
[579]
[580]
[581]
[582]
[583]
[584]
[585]
[586]
[587]
[588]
[589]
[590]
[591]
[592]
[593]
[594]
[595]
[596]
[597]
[598]
[599]
[600]
[601]
[602]
[603]
[604]
[605]
[606]
[607]
[608]
[609]
[610]
[611]
[612]
[613]
[614]
[615]
[616]
[617]
[618]
[619]
[620]
[621]
[622]
[623]
[624]
[625]
[626]
[627]
[628]
[629]
[630]
[631]
[632]
[633]
[634]
[635]
[636]
[637]
[638]
[639]
[640]
[641]
[642]
[643]
[644]
[645]
[646]
[647]
[648]
[649]
[650]
[651]
[652]
[653]
[654]
[655]
[656]
[657]
[658]
[659]
[660]
[661]
[662]
[663]
[664]
[665]
[666]
[667]
[668]
[669]
[670]
[671]
[672]
[673]
[674]
[675]
[676]
[677]
[678]
[679]
[680]
[681]
[682]
[683]
[684]
[685]
[686]
[687]
[688]
[689]<

Complete proteome.

FT	METAL	107	107	MAGNESIUM.
FT	METAL	109	109	MAGNESIUM.
FT	METAL	151	151	MAGNESIUM.
FT	BINDING	158	189	PCP DOMAIN OF SUBSTRATE (POTENTIAL).
FT	VARIANT	157	224	EGKSLPLDSFSLRHODQVSELDSDSPCKIKYEVD (IN NON SURFACTIN-PRODUCING STRAINS).
FT	MUTAGEN	105	105	G->A: ALMOST NO ACTIVITY.
FT	MUTAGEN	105	105	G->D: LOSS OF ACTIVITY.
FT	MUTAGEN	107	107	D->E: 3000-FOLD REDUCTION IN ACTIVITY.
FT	MUTAGEN	147	147	BUT NO CHANGE IN SUBSTRATE AFFINITY.
FT	MUTAGEN	147	147	W->A: 24-FOLD REDUCTION IN ACTIVITY.
FT	MUTAGEN	151	151	BUT NO CHANGE IN SUBSTRATE AFFINITY.
FT	MUTAGEN	155	155	W->F: 5-FOLD REDUCTION IN ACTIVITY.
FT	CONFLICT	97	97	BUT NO CHANGE IN SUBSTRATE AFFINITY.
FT	CONFLICT	118	119	E->A: LOSS OF ACTIVITY.
FT	STRAND	2	7	K->A: LOSS OF ACTIVITY.
FT	HELIX	14	21	T->S (IN REF. 2, 3 AND 5).
FT	HELIX	22	23	G->C (IN REF. 2, 3 AND 5).
FT	HELIX	26	34	IA->MP (IN REF. 4).
FT	HELIX	38	58	
FT	STRAND	59	60	
FT	HELIX	63	65	
FT	STRAND	69	70	
FT	STRAND	72	73	
FT	STRAND	76	77	
FT	STRAND	79	80	
FT	STRAND	82	83	
FT	STRAND	85	91	
FT	STRAND	92	93	
FT	STRAND	94	100	
FT	STRAND	105	110	
FT	HELIX	116	119	
FT	HELIX	125	133	
FT	HELIX	136	157	
FT	STRAND	160	161	
FT	STRAND	169	172	
FT	HELIX	174	181	
FT	STRAND	178	181	
FT	STRAND	185	186	
FT	STRAND	190	194	
FT	STRAND	198	199	
FT	STRAND	200	206	
FT	STRAND	216	217	
FT	STRAND	220	224	
SO	SEQUENCE	224 AA;	26135 MW;	6CBCT6545B8BDF7 CRC64;

Query Match 59.7%; Score 40; DB 1; Length 224;
Best Local Similarity 54.5%; Pred. No. 3.3;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 STDYGRPCLP 12
Db 69 STQYGRPCLP 79

RESULT 7
CPDG_CAVPO STANDARD; PRT; 500 AA.

AC 064403; 054866;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Cytochrome P450 2D16 (EC 1.14.14.1) (CYP2D16).
GN CYP2D16.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Euthera; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-38.
RC STRAIN=13; TISSUE=Adrenal cortex;
RX MEDLINE=95251703; PubMed=7733969;
RA Jiang Q., Voigt J.M., Colby H.D.;
RT "Molecular cloning and sequencing of a guinea pig cytochrome P4502D
RT (CYP2D16): high level expression in adrenal microsomes";
RL Biochem. Biophys. Res. Commun. 209:1149-1156(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley white; TISSUE=Adrenal gland;
RA Sun Y., Voigt J.M., Pierce J.C., Colby H.D.;
RT "The gene sequence of a xenobiotic metabolism-related cytochrome P450
RT isozyme (CYP2D16) in guinea pig adrenal gland";
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE INNER ZONE OF
CC THE ADRENAL CORTEX.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U21486; AAA68479.1; -
CC EMBL; AF020345; AAB94568.1; -
CC PIR; JC4153; JC4153.
CC HSP; P00179; 1DT6.
CC InterPro; IPR001128; Cytochrome_P450.
CC InterPro; IPR008069; BP450_CYP2D.
CC Pfam; PF00067; P450_1.
CC PRINTS; PR01686; BP450CYP2D.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
CC Microsome; Cytochrome P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC Microsome; Endoplasmic reticulum.
FT METAL 81 84 POLY-VAL.
FT DOMAIN 81 84 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 123 123 I->V (IN REF. 2).
FT CONFLICT 127 127 Y->N (IN REF. 2).
FT CONFLICT 148 148 G->R (IN REF. 2).
SQ SEQUENCE 500 AA; 55800 MW; 2429247E49BF6B24 CRC64;
Query Match 59.7%; Score 40; DB 1; Length 500;
Best Local Similarity 72.7%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 STDYGRPCLP 12
Db 98 STDYGRPCLP 108

RESULT 8
UROK_BOVIN STANDARD; PRT; 433 AA.

AC 005589; 028209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

(U-plasminogen activator).

GN PLAV.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; OX NCBI_TaxId=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Aortic endothelium;

RX MEDLINE=93216119; PubMed=8385052;

RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B., Schlemming W.-D.;

RT "Bovine urokinase-type plasminogen activator and its receptor: cloning and induction by retinoic acid.";

RL Gene 125:177-183(1993).

RN [2]

RP SEQUENCE OF 12-433 FROM N.A.

RC TISSUE=Kidney;

RA Ravn P., Berglund L., Petersen T.E.;

RT "Cloning and characterization of the bovine plasminogen activators uPA and tPA.";

RL Inc. Dally J. 5:605-617(1995)

CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

CC -1- INDUCTION: By retinoic acid.

CC -1- SIMILARITY: Belongs to peptidase family S1.

CC -1- SIMILARITY: Contains 1 kringle domain.

CC -1- SIMILARITY: Contains 1 kringle domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L03546; AAA51419.1; -.

DR EMBL; X85801; CA559796.1; -.

DR PIR; JN0560; JN0560.

DR HSSP; P00749; 1LMW.

DR MEROPS; S01.231; -.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR008293; Pept_S1A_UPA.

DR InterPro; IPR001254; Peptidase_S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PIRSF; PIRSF001144; Utk_plasm_act; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; 1.

DR PROSITE; PS01186; EGF_2; FALSE_NEG.

DR PROSITE; PS00026; EGF_3; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Signal; Zymogen.

KW SIGNAL.

FT CHAIN 1 20

FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.

FT CHAIN 181 433 CHAIN A (BY SIMILARITY).

FT CHAIN 181 433 CHAIN B (BY SIMILARITY).

FT DOMAIN 29 65 EGF-LIKE.

FT DOMAIN 72 153 KRINGLE.

FT DOMAIN 154 180 CONNECTING PEPTIDE.

FT DOMAIN 181 433 SERINE PROTEASE.

FT DISULFID 33 41 BY SIMILARITY.

FT DISULFID 35 53 BY SIMILARITY.

FT DISULFID 55 64 BY SIMILARITY.

FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).

FT DISULFID 211 227 BY SIMILARITY.

FT DISULFID 219 290 BY SIMILARITY.

FT DISULFID 315 384 BY SIMILARITY.

FT DISULFID 347 363 BY SIMILARITY.

FT DISULFID 374 402 BY SIMILARITY.

FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT CONFLICT 189 189 A -> T (IN REF. 2).

SO SEQUENCE 433 AA; 48730 MW; 4DE1B8DADA47027A CRC64;

Query Match 58.2%; Score 39; DB 1; Length 433;

Best Local Similarity 63.6%; Pred. No. 10;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ASTDTMGRPCL 11

DB 84 ANRDLGRPCL 94

RESULT 9

YHS4 CAEEL STANDARD; PRT; 602 AA.

ID YHS4 CAEEL

AC 018303;

DT 15-JUL-1999 (Rel. 38, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein ZK849.4 in chromosome 1.

GN ZK849.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

OC Rhabdittidae; Peloderrinae; Caenorhabditis.

OX NCBI_TaxId=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Kershaw J.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP REVISIONS.

RA Durbin R.;

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the Deshrophin family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Z82095; CAB05027.2; -.

DR WormPep; ZK849.4; CB25696.

DR InterPro; IPR000615; Bestrophin.

DR Pfam; PF01062; Bestrophin; 1.

DR ProDom; PD002802; Worm_fam_8; 1.

KW Hypothetical protein.

SO SEQUENCE 602 AA; 67607 MW; 458AB78802BD6B3A CRC64;

Query Match 58.2%; Score 39; DB 1; Length 602;

Best Local Similarity 66.7%; Pred. No. 14;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 DTMGRPCLP 12

DB 416 DTGKDCLP 424

FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 76 76 GAMMA-CARBOXYGLUTAMIC ACID.
 FT DISULFID 61 66 BY SIMILARITY.
 FT DISULFID 91 104 BY SIMILARITY.
 FT DISULFID 109 187 BY SIMILARITY.
 FT DISULFID 130 170 BY SIMILARITY.
 FT DISULFID 158 182 BY SIMILARITY.
 FT DISULFID 215 293 BY SIMILARITY.
 FT DISULFID 236 276 BY SIMILARITY.
 FT DISULFID 264 288 BY SIMILARITY.
 FT DISULFID 333 479 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 338 404 BY SIMILARITY.
 FT DISULFID 533 547 BY SIMILARITY.
 FT DISULFID 561 591 BY SIMILARITY.
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 618 AA; 70268 MW; B89F719A6D601E0 CRC64;

Query Match 58.2%; Score 39; DB 1; Length 618;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 TMGRPLP 12
 Db 231 TLGSPCLP 238

RESULT 11
 PTRR_PIG STANDARD; PRT; 585 AA.

AC P50133; 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Parathyroid hormone/parathyroid hormone-related peptide receptor
 DE Precursor (PTH/PTHr receptor) (PTH/PTHr type I receptor).
 GN PTRR1 OR PTHR.

OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]

SEQUENCE FROM N.A.
 MEDLINE=96305358; PubMed=8688470;
 RA Black E.C., Smith D.P., Zhang X.Y., Frolik C.A., Harvey A.,
 RA Chandrasekhar S., Heilung H.M.,
 RT "Structure and functional expression of a complementary DNA for
 RT porcine parathyroid hormone/parathyroid hormone-related peptide
 RT receptor";
 RL Biochem. Biophys. Acta 1307:339-347(1996).

CC -! FUNCTION: This is a receptor for parathyroid hormone and for
 CC parathyroid hormone-related peptide. The activity of this receptor
 CC is mediated by G proteins which activate adenyl cyclase and also
 CC a phosphatidylinositol-calcium second messenger system (By
 CC similarity).
 CC -! SUBCELLULAR LOCATION: Integral membrane protein.

CC -! SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U18315; AAC48619.1; -.
 DR HSBP; 003431; IBLI.
 DR InterPro; IPR000832; GPCR_secretin.
 DR InterPro; IPR001879; hormn_receptor.

DR Pfam; PF00002; 7tm_2; 1.
 DR Pfam; PF02793; HRM; 1.
 DR PRINTS; P00249; GPCRSECRETIN.
 DR SMART; SM00008; Hormr; 1.
 DR PROSITE; PS00649; G_PROTEIN_REC_P2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_REC_P2_2; 1.
 DR PROSITE; PS50227; G_PROTEIN_REC_P2_3; 1.
 DR PROSITE; PS50261; G_PROTEIN_REC_P2_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 585 PARATHYROID HORMONE/PARATHYROID HORMONE-
 FT RELATED PEPTIDE RECEPTOR.
 FT DOMAIN 27 184 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 185 208 1 (POTENTIAL).
 FT DOMAIN 209 215 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 216 235 2 (POTENTIAL).
 FT DOMAIN 236 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 3 (POTENTIAL).
 FT DOMAIN 302 315 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 316 337 4 (POTENTIAL).
 FT DOMAIN 338 356 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 357 377 5 (POTENTIAL).
 FT DOMAIN 378 404 6 (POTENTIAL).
 FT TRANSMEM 405 423 6 (POTENTIAL).
 FT DOMAIN 424 435 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 436 458 7 (POTENTIAL).
 FT DOMAIN 459 585 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 48 113 BY SIMILARITY.
 FT DISULFID 104 144 BY SIMILARITY.
 FT DISULFID 127 166 BY SIMILARITY.
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 585 AA; 65682 MW; 60BE15CD49B7D210 CRC64;

Query Match 56.7%; Score 38; DB 1; Length 585;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 GRPCLP 12
 Db 101 GRPCLP 106

RESULT 12
 PTRR_MOUSE STANDARD; PRT; 591 AA.

AC P41593; 062119;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Parathyroid hormone/parathyroid hormone-related peptide receptor
 DE Precursor (PTH/PTHr receptor) (PTH/PTHr type I receptor).
 GN PTRR1 OR PTHR.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

SEQUENCE FROM N.A.
 STRAIN=C3H/HEHA;
 MEDLINE=95034305; PubMed=7524627;
 RA Kaptein M., van Dijk T.B., Hoeijmakers T., Cremers F.,
 RA Abou-Samra A.-B., Boonstra J., de Laat S.W., Defize L.H.K.;
 RT "Expression pattern of parathyroid hormone/parathyroid hormone
 RT related peptide receptor mRNA in mouse postimplantation embryos
 RT indicates involvement in multiple developmental processes";
 RL Mech. Dev. 47:29-42(1994).
 RN [2]

DR SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RC MEDLINE=9425468; PubMed=8197183;

DR PROSITE; PSS0261; G PROTEIN RECBP F2 4; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 591
 FT DOMAIN 27 188
 FT TRANSMEM 189 212
 FT DOMAIN 213 219
 FT TRANSMEM 220 239
 FT DOMAIN 240 282
 FT TRANSMEM 283 306
 FT DOMAIN 307 320
 FT TRANSMEM 321 342
 FT DOMAIN 343 361
 FT TRANSMEM 362 382
 FT DOMAIN 383 409
 FT TRANSMEM 410 428
 FT DOMAIN 429 440
 FT TRANSMEM 441 463
 FT DOMAIN 464 591
 FT DISULFID 48 117
 FT DISULFID 108 148
 FT DISULFID 131 170
 FT CARBOHYD 151 151
 FT CARBOHYD 161 161
 FT CARBOHYD 166 166
 FT CARBOHYD 176 176
 SQ SEQUENCE 591 AA; 66260 MW; 21944F3051B9B9C1 CMC64;

Query Match 56.7%; Score 38; DB 1; Length 591;
 Best local similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Gy 7 GRPCLP 12
 Db 105 GRPCLP 110

RESULT 14

PTRR_HUMAN STANDARD; PRT; 593 AA.
 ID_PTRR_HUMAN
 AC Q03431;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Parathyroid hormone/parathyroid hormone-related peptide receptor
 DE precursor (PTH/PTHrP receptor) (PTH/PTHrP type I receptor).
 GN PTHRI OR PTHR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=93238641; PubMed=8386612;
 RA Schipani E., Karga H., Karaplis A.C., Potts J.T. Jr., Kronenberg H.M.,
 RA Abou-Samra A.-B., Segre G.V., Jueppner H.;
 RT "Identical complementary deoxyribonucleic acids encode a human renal
 RT and bone parathyroid hormone (PTH)/PTH-related peptide receptor.";
 RL Endocrinology 132:2157-2165(1993).
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=93387403; PubMed=8397094;
 RA Schneider H., Feyen U.-H., Rao Movva N.;
 RT "Cloning and functional expression of a human parathyroid hormone
 RT receptor.";
 RL Eur. J. Pharmacol. 246:149-155(1993).
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95263723; PubMed=7745008;
 RA Schipani E., Weinstein L.S., Bergwitz C., Iida-Klein A., Kong X.F.,

RA Stuhmann M., Kruse K., Whyte M.P., Murray T., Schmidtke J., Dop C.,
 RA Brickman A.S., Crawford J.D., Potts J.T. Jr., Kronenberg H.M.,
 RA Abou-Samra A.-B., Segre G.V., Jueppner H.;
 RT "Pseudohypoparathyroidism type 1b is not caused by mutations in the
 RT coding exons of the human parathyroid hormone (PTH)/PTH-related
 RT peptide receptor gene.";
 RL J. Clin. Endocrinol. Metab. 80:1611-1621(1995).
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Levine M.A.;
 RT "Characterization of cDNA and genomic DNA encoding the human PTH/PTHrP
 RT receptor.";
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN
 RP EXTRACELLULAR DOMAIN DISULFIDE BONDS.
 RX MEDLINE=20374568; PubMed=10913300;
 RA Graecheop U., Lilje H., Honold K., Wozny M., Reusch D., Baswein A.,
 RA Schaefer W., Ruchnagel K.P., Rudolph R.;
 RT "The N-terminal fragment of human parathyroid hormone receptor 1
 RT constitutes a hormone binding domain and reveals a distinct disulfide
 RT pattern.";
 RL Biochemistry 39:8878-8887(2000).
 RN
 RP STRUCTURE BY NMR OF 168-198.
 RX MEDLINE=98409426; PubMed=9737850;
 RA Pellegrini M., Bisello A., Rosenblatt M., Chorev M., Mierke D.F.;
 RT "Binding domain of human parathyroid hormone receptor: from
 RT conformation to function.";
 RL Biochemistry 37:12737-12743(1998).
 RN
 RP VARIANT JMC ARG-223.
 RX MEDLINE=95215874; PubMed=7701349;
 RA Schipani E., Kruse K., Jueppner H.;
 RT "A constitutively active mutant PTH-PTHrP receptor in Jansen-type
 RT metaphyseal chondrodysplasia.";
 RL Science 268:98-100(1995).
 RN
 RP VARIANTS JMC ARG-223 AND PRO-410.
 RX MEDLINE=9636745; PubMed=8703170;
 RA Schipani E., Langman C.B., Parfitt A.M., Jensen G.S., Kikuchi S.,
 RA Koeh S.W., Cole W.G., Jueppner H.;
 RT "Constitutively activated receptors for parathyroid hormone and
 RT parathyroid hormone-related peptide in Jansen's metaphyseal
 RT chondrodysplasia.";
 RL New Engl. J. Med. 335:708-714(1996).
 RN
 RP CHARACTERIZATION OF VARIANTS JMC ARG-223 AND PRO-410.
 RX MEDLINE=97322091; PubMed=9178745;
 RA Schipani E., Jensen G.S., Pincus J., Nissensohn R.A., Gardella T.J.,
 RA Jueppner H.;
 RT "Constitutive activation of the cyclic adenosine 3',5'-monophosphate
 RT signaling pathway by parathyroid hormone (PTH)/PTH-related peptide
 RT receptors mutated at the two loci for Jansen's metaphyseal
 RT chondrodysplasia.";
 RL Mol. Endocrinol. 11:851-858(1997).
 RN
 RP VARIANT BOCAL LEO-132.
 RX MEDLINE=98417978; PubMed=9745456;
 RA Zhang P., Jobert A.-S., Couvineau A., Silve C.;
 RT "A homozygous inactivating mutation in the parathyroid
 RT hormone/parathyroid hormone-related peptide receptor causing
 RT Blomstrand chondrodysplasia.";
 RL J. Clin. Endocrinol. Metab. 83:3365-3368(1998).
 RN
 RP VARIANT JMC ARG-458.
 RX MEDLINE=99415605; PubMed=10487664;
 RA Schipani E., Langman C.B., Hunzelman J., Le Merrer M., Loke K.Y.,
 RA Dillon M.J., Silve C., Jueppner H.;
 RT "A novel parathyroid hormone (PTH)/PTH-related peptide receptor
 RT mutation in Jansen's metaphyseal chondrodysplasia.";
 RL J. Clin. Endocrinol. Metab. 84:3052-3057(1999).
 RN

RP VARIANT ENCHONDROMATOSIS CYS-150.
 RX MEDLINE=31918583; PubMed=11850620;
 RA Hopyan S., Gokgoz N., Poon R., Gensure R.C., Yu C., Cole W.G.,
 RA Bell R.S., Jueppner H., Andrusis I.L., Wunder J.S., Alman B.A.;
 RT "A mutant PTHrP type I receptor in enchondromatosis.";
 RL Nat. Genet. 30:306-310(2002).
 CC -1- FUNCTION: This is a receptor for parathyroid hormone and for
 CC parathyroid hormone-related peptide. The activity of this
 CC receptor is mediated by G proteins which activate adenylyl
 CC cyclase and also a phosphatidylinositol-calcium second messenger
 CC system.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in most tissues. Most abundant in
 CC kidney, bone and liver.
 CC -1- DISEASE: Defects in PTHRI are the cause of Jansen's metaphyseal
 CC chondrodysplasia (JMC) [MIM:156400]. JMC is a rare autosomal
 CC dominant disorder characterized by a short-limbed dwarfism
 CC associated with hypercalcemia and normal or low serum
 CC concentrations of the two parathyroid hormones.
 CC -1- DISEASE: Defects in PTHRI are the cause of chondrodysplasia
 CC Blomstrand type (BOCD) [MIM:215045]. BOCD is a severe skeletal
 CC dysplasia.
 CC -1- DISEASE: Defects in PTHRI can be a cause of enchondromatosis
 CC [MIM:166000]. Enchondromas are common benign cartilage tumors of
 CC bone. They can occur as solitary lesions or as multiple lesions in
 CC enchondromatosis (Ollier and Maffucci diseases). Clinical problems
 CC caused by enchondromas include skeletal deformity and the
 CC potential for malignant change to osteosarcoma.
 CC -1- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L04308; AAA36525.1; -;
 DR EMBL; X68596; CAA48589.1; -;
 DR EMBL; U22409; AAB60657.1; -;
 DR EMBL; U22401; AAB60657.1; JOINED.
 DR EMBL; U22402; AAB60657.1; JOINED.
 DR EMBL; U22403; AAB60657.1; JOINED.
 DR EMBL; U22404; AAB60657.1; JOINED.
 DR EMBL; U22405; AAB60657.1; JOINED.
 DR EMBL; U22406; AAB60657.1; JOINED.
 DR EMBL; U22407; AAB60657.1; JOINED.
 DR EMBL; U22408; AAB60657.1; JOINED.
 DR EMBL; U17418; AAA56774.1; -;
 DR PIR; I38139; A49191.
 DR PDB; 1BL1; 30-MAR-89.
 DR PDB; 1ET2; 06-SEP-00.
 DR PDB; 1ET3; 06-SEP-00.
 DR Genew; HGNC:9608; PTHRI.
 DR MIM; 168468; -;
 DR MIM; 156400; -;
 DR MIM; 215045; -;
 DR MIM; 166000; -;
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0004991; F:parathyroid hormone receptor activity; TAS.
 DR GO; GO:0007187; P:G-protein signaling, coupled to cyclic nucl. . .; TAS.
 DR GO; GO:0001501; P:skeletal development; TAS.
 DR InterPro; IPR000832; GPCR_secretin.
 DR InterPro; IPR001879; horm_receptor.
 DR Pfam; PFO0002; 7tm_2; 1.
 DR Pfam; PF02793; HRM; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR SMART; SMD0008; Hormr; 1.
 DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.

DR PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
 DR PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Disease mutation; 3D-structure; Dwarfism.
 FT SIGNAL 1 26
 FT CHAIN 27 593
 FT DOMAIN 27 188
 FT TRANSMEM 189 212
 FT DOMAIN 213 219
 FT TRANSMEM 220 239
 FT DOMAIN 240 282
 FT TRANSMEM 283 306
 FT DOMAIN 307 320
 FT TRANSMEM 321 342
 FT DOMAIN 343 361
 FT TRANSMEM 362 382
 FT DOMAIN 383 409
 FT TRANSMEM 410 428
 FT DOMAIN 429 440
 FT TRANSMEM 441 463
 Query Match 56.7%; Score 38; DB 1; Length 593;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 7 GRPCLP 12
 Db 105 GRPCLP 110
 RESULT 15
 PSB3 TRYBB STANDARD; PRT; 205 AA.
 ID PSB3 TRYBB
 AC QGNDAI;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Proteasome subunit beta type 3 (EC 3.4.25.1) (20S proteasome subunit
 DE beta-3).
 GN PSB3.
 OS Trypanosoma brucei brucei.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=427;
 RA Bozdech Z., Huang L., Morton A., Wang C.C.;
 RT "bPpsb3-beta 3 subunit of 20S proteasome from T. brucei.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: The proteasome is a multicatalytic proteinase complex
 CC which is characterized by its ability to cleave peptides with Arg,
 CC Phe, Tyr, Leu, and Glu adjacent to the leaving group at neutral or
 CC slightly basic pH. The proteasome has an ATP-dependent proteolytic
 CC activity (By similarity).
 CC -1- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
 CC specificity.
 CC -1- PATHWAY: Involved in an ATP/ubiquitin-dependent non-lysosomal
 CC proteolytic pathway.
 CC -1- SUBUNIT: The proteasome is composed of at least 15 non identical
 CC subunits which form a highly ordered ring-shaped structure (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family 71B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF169653; AAP89685.1; -.
 DR HSSP; P25451; IRYP.
 DR InterPro; IPR000243; Pept_T1A_subB.
 DR InterPro; IPR001353; Peptidase_T1.
 DR Pfam; PF00227; Proteasome; 1.
 DR PROSITE; PS00854; PROTEASOME_B; FALSE_NEG.
 KW Proteasome; Hydrolase; Protease; Threonine protease.
 SQ SEQUENCE 205 AA; 22458 MW; 4CB209381E2FFD7 CRC64;

Query Match 55.2%; Score 37; DB 1; Length 205;
 Best Local Similarity 54.5%; Pred. No. 11;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 STDTPGRPCLP 12
 :||:|
 Db 123 ATDLIGACEP 133

Search completed: March 18, 2004, 11:17:01
 Job time : 1.73523 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:09:25 ; Search time 3.28228 seconds
(without alignments)
1153.535 Million cell updates/sec

Title: US-10-076-421-4
Perfect score: 67
Sequence: 1 ASTDTWGRPLP 12

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteria:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	100.0	154	4 Q96SE8	Q96SE8 homo sapien
2	51	76.1	231	11 Q8C6L2	Q8C6L2 mus musculi
3	44	65.7	433	6 Q8M1L0	Q8M1L0 oryctolagus
4	44	65.7	433	6 Q8M1L0	Q8M1L0 oryctolagus
5	42	62.7	402	17 Q8TYC9	Q8TYC9 methanopyru
6	40	59.7	165	2 P94401	P94401 bacillus su
7	40	59.7	165	2 Q93UT8	Q93UT8 bacillus su
8	40	59.7	253	2 Q7WK66	Q7WK66 alcaligenes
9	40	59.7	581	16 Q87SD8	Q87SD8 vibrio para
10	40	59.7	608	16 Q8BZW7	Q8BZW7 leprospira
11	40	59.7	787	3 Q96U20	Q96U20 neurospora
12	39	58.2	27	10 Q9S8W4	Q9S8W4 avena sativ
13	39	58.2	157	6 Q9TV48	Q9TV48 bos taurus
14	39	58.2	288	3 Q14136	Q14136 schizosach
15	39	58.2	419	5 Q9VB45	Q9VB45 drosophila
16	39	58.2	419	5 Q24548	Q24548 drosophila

17	39	58.2	434	10 Q7XQ04	Q7XQ04 oryza sativ
18	39	58.2	476	3 Q8NJK6	Q8NJK6 aspergillus
19	39	58.2	477	3 Q93J36	Q93J36 streptomyce
20	39	58.2	533	16 Q82NM7	Q82NM7 streptomyce
21	39	58.2	543	16 Q8YY82	Q8YY82 anabaena sp
22	39	58.2	621	10 Q7X619	Q7X619 oryza sativ
23	39	58.2	800	10 Q8S868	Q8S868 oryza sativ
24	39	58.2	800	10 Q7XEL6	Q7XEL6 oryza sativ
25	39	58.2	813	5 Q9BLJ2	Q9BLJ2 ciona intes
26	39	58.2	1063	5 Q8SR27	Q8SR27 encephalito
27	38	56.7	102	16 Q7UMM5	Q7UMM5 rhodospirill
28	38	56.7	125	10 Q9SUZ2	Q9SUZ2 arabidopsis
29	38	56.7	167	17 Q9YAM3	Q9YAM3 aeropyrum p
30	38	56.7	232	16 Q8CUZ6	Q8CUZ6 oceanobacil
31	38	56.7	259	10 Q9ARS4	Q9ARS4 oryza sativ
32	38	56.7	333	4 Q8NSV1	Q8NSV1 homo sapien
33	38	56.7	437	9 Q9AZM2	Q9AZM2 bacterioph
34	38	56.7	437	16 Q9C194	Q9C194 lactococcus
35	38	56.7	540	16 Q9K414	Q9K414 streptomyce
36	38	56.7	589	6 Q9GMD1	Q9GMD1 oryctolagus
37	38	56.7	589	6 Q7YR13	Q7YR13 cervus elap
38	38	56.7	591	11 Q91WV4	Q91WV4 mus musculi
39	38	56.7	591	11 Q80WU8	Q80WU8 mus musculi
40	38	56.7	595	6 Q9TU31	Q9TU31 canis famli
41	38	56.7	618	11 Q99KV8	Q99KV8 mus musculi
42	38	56.7	668	11 Q8R5A3	Q8R5A3 mus musculi
43	38	56.7	688	3 Q87IU3	Q87IU3 neurospora
44	38	56.7	877	10 Q7XU03	Q7XU03 oryza sativ
45	38	56.7	3956	2 Q7WTF2	Q7WTF2 streptomyce

ALIGNMENTS

RESULT 1
ID Q96SE8 PRELIMINARY; PRT; 154 AA.
AC Q96SE8;
DT 01-DEC-2001 (TREMREL. 19, Created)
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL: AY029537; AK38734.1; -;
DR GO: GO:0016301; F-kinase activity; IEA.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; Kringle; 1.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.

```

DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CF2FCFF505572 CRC64;

Query Match 100.0%; Score 67; DB 4; Length 154;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCLP 12
Db 82 ASTDTMGRPCLP 93

RESULT 2
ID Q86GL2 PRELIMINARY; PRT; 231 AA.
AC Q86GL2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA MEDLINE=22354683; PubMed=12466851;
RX THE FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC3743.1; -.
DR PIR; P70534; P70534.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
FT NON TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25F8980A682737F2 CRC64;

Query Match 76.1%; Score 51; DB 11; Length 231;
Best Local Similarity 81.8%; Pred. No. 0.23;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCL 11
Db 83 ANTDTMGRPCL 93

RESULT 3
ID Q8MTL0 PRELIMINARY; PRT; 433 AA.
AC Q8MTL0;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
OS PLANU.
GN Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCB1_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2215945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY122285; AA83187.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR008293; Pept_S1A_uPA.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR PIRSF; PIRSF001144; Urk plas act; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65B64F3641554980 CRC64;

Query Match 65.7%; Score 44; DB 6; Length 433;
Best Local Similarity 72.7%; Pred. No. 8.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ASTDTMGRPCL 11
Db 84 ANTDTMGRPCL 94

RESULT 4
ID Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator.
OS UROKINASE.
GN Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCB1_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator. mRNA,
RT complete cds."

```


	Query Match	65.7%	Score 44;	DB 6;	Length 433;
	Best Local Similarity	72.7%	Pred. NO. 8.4;		
Matches	8;	Conservative	1;	Mismatches	2; Indels 0; Gaps 0.
Db	1 ASTDTMRPCL 11 : 84 ANTDIMDRPCL 94				
RESULT 5					
Q8TYC9	PRELIMINARY;	PRT;	402 AA.		
AC Q8TYC9;					
DT 01-JUN-2002 (TREMBlrel. 21. Created)					
DT 01-JUN-2002 (TREMBlrel. 21. Last sequence update)					
DT 01-JUN-2003 (TREMBlrel. 24. Last annotation update)					
DE Predicted GPase, probable translation factor.					
GN MK0373.					
OS Methanopyrus kandleri.					
OC Archaeae; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;					
CC Methanopyrus					
CX NCBI_TaxID=2320;					
RN [1]					
RP SEQUENCE FROM N.A.					
RA STRAIN=AV19 / DSM 6324 / JCM 9639;					
RX MEDLINE=11927647; PubMed=11930014;					
SL Slegarev A.I., Mezhevaeva K.V., Makarova K.S., Polishin N.N., Shcherbdina O.V., Shaknova V.V., Belova G.I., Aravind L., Natalek D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O., Malysk A.G., Koonin E.V., Kozlyavsk S.A.; "The complete genome of hyperthermophile Methanopyrus kandleri AV19 RT and monophyly of archaeal methanogens." Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).					
RL EMBL; AB010334; AAC01588.1; -.					
DR InterPro; IPRO06073; GTP1_OBG.					
DR InterPro; IPRO04095; TGS_dom.					
DR Pfam; PF02824; TGS; 1.					
DR PRINTS; PR00326; GTP1OBG.					
KM Complete proteome.					
SO SEQUENCE 402 AA; 44508 MW; 6DVSJ31395SD736894 CRC64;					

OY	2	STDTMGRPCLP	12
Db	116	STDEGRPCDP	126
RESULT 6			
ID	P94401	PRELIMINARY;	PRT; 165 AA.
AC	P94401;		
DT	01-MAY-1997 (TREMBlrel. 03. Created)		
DT	01-MAY-1987 (TREMBlrel. 03. Last sequence update)		
DE	01-OCT-2003 (TREMBlrel. 25. Last annotation update)		
GN	Sufactin production.		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
NB	NCBI_TaxID=1423;		
XX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=168 trpC2.		
RX	MEDLINE=97124189; PubMed=8969502;		
RA	Yamane K., Kumano M., Kurita K.;		
RT	"The 25 degree-36 degree region of the Bacillus subtilis chromosome:		
RT	determination of the sequence of a 146kb segment and identification of		
RT	113 genes."		
RL	Microbiology 142:3047-3056(1996).		
DR	EMBL; D50453; BAA08991.1; -.		
DR	HSSP; P39135; 1080.		
DR	InterPro; IPR008278; 4-PPT_transf.		
DR	Pfam; PF01648; ACPS; 1.		
SO	SEQUENCE 165 AA; 19580 MW; OCSES001AFD465DF CRC64;		
Query Match 59.7%; Score 40; DB 2; Length 165;			
Best Local Similarity 54.5%; Pred. No. 18;			
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;			
OY	2	STDTMGRPCLP	12
Db	69	STGEYKRPCIP	79
RESULT 7			
ID	O93UT8	PRELIMINARY;	PRT; 165 AA.
AC	O93UT8;		
DT	01-DEC-2001 (TREMBlrel. 19. Created)		
DT	01-DEC-2001 (TREMBlrel. 19. Last sequence update)		
DE	01-OCT-2003 (TREMBlrel. 25. Last annotation update)		
GN	Inactive surfactin production protein.		
OS	Sfp-0.		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
NB	NCBI_TaxID=1423;		
XX	[1]		
RP	SEQUENCE FROM N.A.		
RC	Roonagawang N., Morikawa M., Kanaya S.;		
RT	"Sfp-0 gene of a surfactin nonproducing Bacillus subtilis Mill3."		
RL	Submitted (May-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB061356; BAB47573.1; -.		
DR	InterPro; IPR008278; 4-PPT_transf.		
DR	Pfam; PF01648; ACPS; 1.		
SO	SEQUENCE 165 AA; 19568 MW; 6BAE919A917B1CA CRC64;		
Query Match 59.7%; Score 40; DB 2; Length 165;			
Best Local Similarity 54.5%; Pred. No. 18;			
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;			
OY	2	STDTMGRPCLP	12

Db 69 STQXGKPCIP 79

RESULT 8

Q7WX66 PRELIMINARY; PRT; 253 AA.
 AC Q7WX66;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Transcriptional regulator.
 GN FNR OR PHG276.
 OS Alkaligenes eutrophus (Ralstonia eutropha).
 OC Plasmid megaplasmid PHG.
 CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 CC Burkholderiaceae; Ralstonia.
 CX NCBI_TaxID=510;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H16;
 RA Schwartz E., Henne A., Cramm R., Eitinger T., Friedrich B.,
 RA Gottschalk G.;
 RT "Complete Nucleotide Sequence of pHG: A Ralstonia eutropha H16
 RT Megaplasmid Encoding Key Enzymes of H2-based Lithoautotrophy and
 RT Anaerobiosis.";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 KW EMBL; AY305378; AAP86025.1; -.
 KM Plasmid.
 SQ SEQUENCE 253 AA; 28316 MW; F664DD75198CE1E9 CRC64;

Query Match 59.7%; Score 40; DB 2; Length 253;
 Best Local Similarity 72.7%; Pred. No. 27;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 STDIMGRCPLP 12
 |||:|
 27 STCTMGQLCLP 37

RESULT 9

Q87SD8 PRELIMINARY; PRT; 581 AA.
 ID Q87SD8;
 AC Q87SD8;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Sensory box/GDEF family protein.
 GN VP0486.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 CC Vibrionaceae; Vibrrio.
 CX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Matino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Nishijima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
 RT distinct from that of V. cholerae.";
 RL Lancet 361:743-749(2003).
 DR EMBL; AP005074; BACS8749.1; -.
 DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
 DR InterPro; IPR001610; GDEF.
 DR InterPro; IPR000700; PAS-aassoc_C.
 DR Pfam; PF00990; GDEF. 1.
 DR Pfam; PF00785; PAC. 1.
 DR PROSITE; PS50887; GDEF. 1.
 DR PROSITE; PS50113; PAC. 2.

KW Complete proteome.
 SQ SEQUENCE 581 AA; 66107 MW; 1830D37B51E24994 CRC64;

Query Match 59.7%; Score 40; DB 16; Length 581;
 Best Local Similarity 70.0%; Pred. No. 61;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 STDIMGRCPL 11
 :|||:|
 Db 216 TTDTQGRSCL 225

RESULT 10

Q8EZW7 PRELIMINARY; PRT; 608 AA.
 ID Q8EZW7;
 AC Q8EZW7;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Conserved hypothetical protein.
 GN LA3734.
 OS Leptospira interrogans.
 CC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 CX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE011529; AAN50932.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 608 AA; 63971 MW; DA173AB9FC21B1F CRC64;

Query Match 59.7%; Score 40; DB 16; Length 608;
 Best Local Similarity 58.3%; Pred. No. 64;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ASTDTMGRCPLP 12
 |||:|
 Db 589 AATDILGRPLP 600

RESULT 11

Q96U20 PRELIMINARY; PRT; 787 AA.
 ID Q96U20;
 AC Q96U20;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN B208.430.
 OS Neurospora crassa.
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 CX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohnselt J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL555930; CAD01134.1; -.
 DR GO; GO:0003743; P:translation initiation factor activity; IEA.
 DR GO; GO:0006413; P:translational initiation; IEA.
 DR InterPro; IPR001950; TIF_SUI1.
 DR PROSITE; PS01118; SUI1_1; 1.
 KW Hypothetical protein
 SQ SEQUENCE 787 AA; 84646 MW; BCCA1937ABEB5B3 CRC64;

Query Match 59.7%; Score 40; DB 3; Length 787;

Best Local Similarity 77.8%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASTDTWGRP 9
Db 219 SSTDSMGRP 227

RESULT 12

O9S8M4 PRELIMINARY; PRT; 27 AA.
AC O9S8M4; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Alpha-amylase inhibitor (Fragment).
OS Avena sativa (Oat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Avenae; Avena.
OX NCBI_TaxID=4498;
RN [1]
RP SEQUENCE
RX MEDLINE=92405739; Pubmed=1526282;
RA Rocher A., Collila F., Ortiz M.L., Mendez E.;
RT "Identification of the three major coeliac immunoreactive proteins and
RL one alpha-amylase inhibitor from oat endosperm."
RL PIR, S29211, S29211.
DR HSP, P07597, 1BE2.
SQ SEQUENCE 27 AA; 2756 MW; E48BD7DA87385341 CRC64;

Query Match 58.2%; Score 39; DB 10; Length 27;
Best Local Similarity 66.7%; Pred. No. 4.5;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 DTWGRPCLP 12
Db 7 DTLVKRCLP 15

RESULT 13

O9TVAB PRELIMINARY; PRT; 157 AA.
AC O9TVAB; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; Pubmed=11204721;
RA Balcerzak D., Querengesser U., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle."
RL J. Anim. Sci. 79:94-107(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL, AF144761; AAD30301.1; -.
DR HSP, P00749, 1URK.
DR GO, GO:0016301, F:kinase activity; IEA.
DR InterPro, IPR006209; EGF_1ike.
DR InterPro, IPR000001; Kringle.
DR Pfam, PF00051; kringle.1.
DR PRINTS, PRO0018; KRINGLE.
DR PRODOM, PD000395; Kringle.1.
DR SMART, SM00130; KR, 1.
DR PROSITE, PS00022; EGF_1; 1.

DR PROSITE, PS00021; KRINGLE 1; 1.
DR PROSITE, PS50070; KRINGLE 2; 1.
KM Glycoprotein; Kinase; Kringle.
FT NON TER 1
FT NON TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBB7.CRC64;

Query Match 58.2%; Score 39; DB 6; Length 157;
Best Local Similarity 63.6%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ASTDTWGRP 11
Db 50 ANRDLGRPCL 60

RESULT 14

O14136 PRELIMINARY; PRT; 288 AA.
AC O14136; 01-JAN-1999 (TREMBlrel. 09, Created)
DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Putative peroxisomal membrane protein, with SRC homology domain.
GN SPAC3C7.10.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycaceae.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Brown D., Churcher C.M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Wood V., Barrall B.G., Rajandream M.A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL, Z99568; CAB16740.1; -.
DR PIR, T38696; T38696.
DR GenDB SPombe; SPAC3C7.10; -.
DR InterPro, IPR007223; Peroxin-13_N.
DR InterPro, IPR001452; SH3.
DR Pfam, PF04088; Peroxin-13_N.1.
DR Pfam, PF00018; SH3.1.
DR PRINTS, PRO0452; SH3DOMAIN.
DR PRODOM, PD000066; SH3.1.
DR SMART, SM00326; SH3.1.
DR PROSITE, PS50002; SH3.1.
KW SH3 domain.
SQ SEQUENCE 288 AA; 32054 MW; 992P93538CDE1E43 CRC64;

Query Match 58.2%; Score 39; DB 3; Length 288;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ASTDTWGRP 10
Db 254 SKDTQGNPC 263

RESULT 15

O9VBAS PRELIMINARY; PRT; 419 AA.
AC O9VBAS; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE T48 protein (GMI8993p).
GN T48 OR CG5507.

OS *Drosophila melanogaster* (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OC NCBI_taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutcliffe G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Baou A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu L.B., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell U.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Fries E.,
 RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Paclet J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celinker S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB003758; AAF56637.1; -;
 DR EMBL: AY118853; AAM50713.1; -;
 DR FlyBase; FBgn0004359; T48.
 DR InterPro; IPR009030; Grow_fac_recip.
 DR SQ SEQUENCE 419 AA; 46210 MW; 213807918BA8918 CRC64;

Query Match 58.2%; Score 39; DB 5; Length 419;
 Best Local Similarity 58.3%; Pred. No. 68;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ASTDTMGRCPLP 12
 DB 135 AHTEFMGRVCVP 146

Search completed: March 18, 2004, 11:19:21
 Job time : 5.28228 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:53:24 ; Search time 4.13567 Seconds
(without alignments)
683.197 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPWCYQVQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	10	5	AAG79463
2	62	100.0	411	2	AAV39343 Human pro
3	62	100.0	411	2	AAV42284 Human pro
4	62	100.0	411	5	AAE13269 Human pro
5	62	100.0	411	7	ADC72159 Human pro
6	57	91.9	65	1	AAV93232 Modified
7	57	91.9	86	7	ABR42617 Human pro
8	57	91.9	86	7	ABR42602 Mouse
9	57	91.9	86	7	ABR42599 Human pro
10	57	91.9	86	7	ABR42601 Human pro
11	57	91.9	87	7	ABR42604 Human pro
12	57	91.9	87	7	ABR42605 Human pro
13	57	91.9	88	5	AAE16542 Human pro
14	57	91.9	89	2	AAW22742 Urokinase
15	57	91.9	96	5	AAE16550 Human pro
16	57	91.9	135	5	AAE16545 Human pro
17	57	91.9	143	5	AAE16549 Human pro
18	57	91.9	194	5	AAW22745 Metastasi
19	57	91.9	201	2	AAW22746 Metastasi
20	57	91.9	322	7	ABR42615 Fusion pr
21	57	91.9	322	7	ABR42616 Fusion pr
22	57	91.9	337	4	AAV75492 Human col
23	57	91.9	337	5	ABP41795 Human ova
24	57	91.9	365	2	AAE68854 Delta 1-4
25	57	91.9	378	2	AAW13635 Human pro

26	57	91.9	386	2	AAE66266	AAE66266 Bifunctio
27	57	91.9	389	2	AAW13636	AAW13636 Human pro
28	57	91.9	390	2	AAE66245	AAE66245 Bifunctio
29	57	91.9	390	2	AAE66247	AAE66247 Bifunctio
30	57	91.9	392	2	AAE66260	AAE66260 Bifunctio
31	57	91.9	392	2	AAE66264	AAE66264 Bifunctio
32	57	91.9	392	2	AAE66255	AAE66255 Bifunctio
33	57	91.9	392	2	AAE66259	AAE66259 Bifunctio
34	57	91.9	392	2	AAE66261	AAE66261 Bifunctio
35	57	91.9	392	2	AAE66258	AAE66258 Bifunctio
36	57	91.9	392	2	AAE66263	AAE66263 Bifunctio
37	57	91.9	392	2	AAE66254	AAE66254 Bifunctio
38	57	91.9	392	2	AAE66256	AAE66256 Bifunctio
39	57	91.9	392	2	AAE66257	AAE66257 Bifunctio
40	57	91.9	393	2	AAE66251	AAE66251 Bifunctio
41	57	91.9	393	2	AAE66253	AAE66253 Bifunctio
42	57	91.9	393	2	AAE66249	AAE66249 Bifunctio
43	57	91.9	393	2	AAE66252	AAE66252 Bifunctio
44	57	91.9	393	2	AAE66244	AAE66244 Bifunctio
45	57	91.9	393	2	AAE66250	AAE66250 Bifunctio

ALIGNMENTS

RESULT 1
AAG79463
ID AAG79463 standard; peptide; 10 AA.
XX
AC AAG79463;
XX
DT 15-NOV-2002 (first entry)
XX
DE sc-uPA long A chain peptide fragment 3, amino acids 129-138.
XX
KW Single chain prepro-urokinase; sc-uPA; pro-urokinase; HIV;
KW high molecular weight urokinase-type plasminogen activator; HMW-uPA;
KW long A; long B; EGF-like domain; kringle domain; urokinase receptor;
KW low molecular weight urokinase-type plasminogen activator; LMW-uPA; CD87;
KW binding domain.
XX
OS Homo sapiens.
XX
PN EPI232755-A2.
XX
PD 21-AUG-2002.
XX
PF 15-FEB-2002; 2002EP-00003555.
XX
PR 20-FEB-2001; 2001JP-00042655.
PR 19-JUN-2001; 2001JP-00184284.
XX
PA (JCRP-) JCR PHARM CO LTD.
XX
PI Wada M, Wada N;
XX
XX WPI; 2002-610512/66.
DR
XX
XX Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g. high
PT molecular weight urokinase-type plasminogen activator, amino-terminal
PT fragment or an anti-CD87 antibody.
XX
XX Example; Page 23; 38pp; English.
XX
XX The sequences given in AAG79461-63 represent peptide fragments derived
XX from the long A chain of single chain prepro-urokinase (sc-uPA). Pro-
XX urokinase (amino acids 21-431) with a cleavage between amino acids 178
XX and 179 gives high molecular weight urokinase-type plasminogen activator
XX (HMW-uPA). HMW-uPA is a protein consisting of two peptide chains linked
XX by a di-sulphide bond. The chains, long A and B, are formed by enzymatic
XX cleavage between amino acids 178 and 179 of pro-urokinase. HMW-uPA
XX includes an EGF-like domain, a kringle domain and a urokinase receptor
XX (CD87) binding domain. HMW-uPA is then cleaved between amino acids 155

CC and 156 to give low molecular weight urokinase-type plasminogen activator
 CC (LMW-UPA) (amino acids 156-178 and 179-431), that has no plasminogen
 CC activator activity. sc-UPA, or these fragments of it, may be used in the
 CC anti-HIV agents of the invention which comprise a ligand molecule that
 CC binds to CD87. The agents are useful for treating HIV-infected humans for
 CC suppression of reproduction of HIV. The anti-HIV agents act by a
 CC mechanism of action different from those of conventional drugs, widening
 CC the choice of therapeutics agents and avoiding problems of resistant HIV
 XX

SO Sequence 10 AA;

Query Match 100.0%; Score 62; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0072;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQVQ 10
 Db 1 RRPWCYQVQ 10

RESULT 2

AA39343

AC AAY39343;

DT 01-DEC-1999 (first entry)

XX Human pro-urokinase.

KW Serine protease; plasminogen; plasmin; activation; matrix; cancer;
 KM tumour; metastasis; X-ray crystallography; inhibitor.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Protein 1. .158

FT Protein /label= Mature_urokinase_A_chain

FT Protein 1. .135

FT Protein /label= Low_molecular_weight_urokinase

FT Disulfide-bond 11. .19

FT Disulfide-bond 13. .31

FT Disulfide-bond 33. .42

FT Disulfide-bond 50. .131

FT Disulfide-bond 71. .113

FT Disulfide-bond 102. .126

FT Cleavage-site 135. .136

FT /note= "Additional cleavage at this site generates low

FT molecular weight (LMW) urokinase"

FT Disulfide-bond 148. .279

FT /note= "Links mature urokinase A- and B-chains"

FT Cleavage-site 158. .159

FT /note= "Cleavage at this site generates mature urokinase

FT Protein A- and B-chains"

FT Protein 159. .411

FT /label= Mature_urokinase_B_chain

FT Disulfide-bond 189. .268

FT Disulfide-bond 197. .268

FT Disulfide-bond 293. .362

FT Modified-site 302

FT /note= "N-glycosylated"

FT Disulfide-bond 325. .341

FT Disulfide-bond 352. .380

FT Cleavage-site 405. .406

XX WO945379-A2.

XX 10-SEP-1999.

XX 05-MAR-1999; 99WO-US004967.

XX 06-MAR-1998; 98US-00036184.

XX

PA (ABBO) ABBOTT LAB.

XX Member VL, Greer J, Abad-Zapatero C, Norbeck DW;

XX WPI; 1999-571607/48.

XX Identifying ligands for target biomolecules using X-ray crystallography.

XX Example 1; Fig 5; 57pp; English.

PS

XX This sequence represents human pro-urokinase. The mature urokinase

CC consists of an A- and B-chain, linked by a single disulfide bond, and is

CC generated by proteolytic cleavage of the peptide bond between Lys 158 and

CC Ile 159. Additional cleavage of the peptide bond between Lys 135 and Lys

CC 136 generates a low molecular weight urokinase. The urokinase A-chain

CC contains an EGF-like domain and a kringle domain, while the B-chain

CC contains the catalytic domain. Urokinase is a serine protease and is

CC strongly associated with tumour cells. Urokinase activates plasminogen

CC which, in turn, activates the matrix metalloproteinases. Plasmin and the

CC metalloproteinases degrade the extracellular matrix and promote tumour

CC growth and metastasis. Inhibitors that specifically target urokinase may

CC serve as effective anticancer agents. A novel method for identifying such

CC ligands used X-ray crystallography to determine if a complex is formed

CC between a ligand and a target biomolecule. However, crystals of a native

CC urokinase/inhibitor complex had poor diffraction quality. Human urokinase

CC was therefore engineered so that it would produce crystals with the

CC desired qualities. This engineered urokinase was designated mu-UK

XX (AAY39344)

XX Sequence 411 AA;

SO

Query Match 100.0%; Score 62; DB 2; Length 411;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQVQ 10

Db 109 RRPWCYQVQ 118

RESULT 3

AA42284

AC AAY42284;

DT 01-DEC-1999 (first entry)

XX Human pro-urokinase.

KW Serine protease; plasminogen; plasmin; activation; matrix; cancer;
 KM tumour; metastasis; X-ray crystallography; inhibitor.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Protein 1. .158

FT Protein /label= Mature_urokinase_A_chain

FT Protein 1. .135

FT Protein /label= Low_molecular_weight_urokinase

FT Disulfide-bond 11. .19

FT Disulfide-bond 13. .31

FT Disulfide-bond 33. .42

FT Disulfide-bond 50. .131

FT Disulfide-bond 71. .113

FT Disulfide-bond 102. .126

FT Cleavage-site 135. .136

FT /note= "Additional cleavage at this site generates low

FT molecular weight (LMW) urokinase"

FT Disulfide-bond 148. .279

FT /note= "Links mature urokinase A- and B-chains"

FT Cleavage-site 158. .159

FT /note= "Cleavage at this site generates mature urokinase

FT		A- and B-chains"
PT	Protein	159. .411
FT	/label= Mature_urokinase_B_chain	
FT	Disulfide-bond	189. .205
FT	Disulfide-bond	197. .268
FT	Disulfide-bond	293. .362
FT	Modified-site	302
FT	/note= "N-glycosylated"	
FT	Disulfide-bond	325. .341
FT	Disulfide-bond	352. .380
XX	Cleavage-site	405. .406
PN	MO9945389-A2.	
PD	10-SEP-1999.	
XX		
PF	01-MAR-1999;	99WO-US004518.
PR	06-MAR-1998;	98US-00036184.
XX	(ABBQ) ABBOTT LAB.	
PA	Nienaber VL, Greer J, Abad-Zapatero C, Norbeck DW;	
P1	WPI, 1999-551079/46.	
DR		
PT	Identifying ligands for target biomolecules using X-ray crystallography,	
PT	used for designing ligands with improved biological activity for target	
PT	receptor.	
PS	Example 1; Fig 5; 57bp; English.	
XX		
CC	This sequence represents human pro-urokinase. The mature urokinase	
CC	consists of an A- and B-chain, linked by a single disulphide bond, and is	
CC	generated by proteolytic cleavage of the peptide bond between Lys 158 and	
CC	Lys 159. Additional cleavage of the peptide bond between Lys 135 and Lys	
CC	136 generates a low molecular weight urokinase. The urokinase A-chain	
CC	contains an EGF-like domain and a kinkle domain, while the B-chain	
CC	contains the catalytic domain. Urokinase is a serine protease and is	
CC	strongly associated with tumour cells. Urokinase activates plasminogen	
CC	which, in turn, activates the matrix metalloproteinases. Plasmin and the	
CC	metalloproteinases degrade the extracellular matrix and promote tumour	
CC	growth and metastasis. Inhibitors that specifically target urokinase may	
CC	serve as effective anticancer agents. A novel method for identifying such	
CC	ligands used X-ray crystallography to determine if a complex is formed	
CC	between a ligand and a target biomolecule. However, crystals of a native	
CC	urokinase/inhibitor complex had poor diffraction quality. Human urokinase	
CC	was therefore engineered so that it would produce crystals with the	
CC	desired qualities. This engineered urokinase was designated mu-UK	
CC	(AAV42285)	
XX		
SQ	Sequence 411 AA;	
Query Match	100.0%; Score 62; DB 2; Length 411;	
Best Local Similarity	100.0%; Pred. No. 0.19;	
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
QY	1 RRPWCYGVQVQ 10	
DB	109 RRPWCYGVQVQ 118	
RESULT 4		
ID	AAE13269 standard; protein; 411 AA.	
AC	AAE13269;	
DT	12-FEB-2002 (first entry)	
XX	Human pro-urokinase (UK) protein.	
DE	Human; urokinase; UK; serine protease; x-ray diffraction;	
XX		

Key	Location/Qualifiers
Domain	/note="A-chain containing EGF-like domain and kringle domain"
Disulfide-bond	11..19
Disulfide-bond	13..31
Disulfide-bond	33..42
Disulfide-bond	50..131
Disulfide-bond	71..113
Disulfide-bond	102..136
Cleavage-site	135..136
Cleavage-site	/label= Proteolytic_cleavage_site
Cleavage-site	148..279
Cleavage-site	158..159
Cleavage-site	/label= Proteolytic_cleavage_site
Domain	159..411
Domain	/note="B-chain containing catalytic serine protease domain"
Disulfide-bond	189..205
Disulfide-bond	197..268
Disulfide-bond	293..361
Disulfide-bond	325..341
Disulfide-bond	352..380
US6297021-B1.	
02-OCT-2001.	
05-MAR-1999;	99US-00263904.
06-MAR-1998;	98US-00036184.
(ABBOTT LAB.	
Nienaber VL, Greer J, Abad-Zapatero C, Norbeck DW;	
WPI; 2002-009432/01.	
Identifying ligand that binds to target biomolecule and designing ligands with improved biological activity for the target receptor, comprises employing x-ray crystallography.	
Example 1; Col 11-12; 33pp; English.	
The invention relates to a method for identifying a ligand that binds to a target biomolecule such as human urokinase (UK) which is a serine protease strongly associated with tumour cells. The method comprises obtaining an x-ray diffraction pattern of the target biomolecule crystal with exposure to a mixture of at least 2 potential ligands. The method is useful for identifying and screening ligands that bind target receptor molecules, and for designing high-affinity ligands (or drugs) with improved biological activity for the target receptor. The present sequence is human pro-urokinase (UK) protein related to the invention	
Sequence 411 AA;	
Query Match	100.0%; Score 62; DB 5; Length 411;
Best Local Similarity	100.0%; Prid. No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
1 RRPNCYQVOQ 10	
109 RRPNCYQVOQ 118	

```

AC  ADC72159;
XX
DT  18-DEC-2003 (first entry)
XX
DE  Human urokinase mature protein sequence related to ligand identification.
XX
XX  ligand; target biomolecule; crystal; X-ray crystal diffraction pattern;
XX  structural change; binding event; crystallographic data;
XX  three dimensional electron density; ligand-receptor complex; human,
XX  urokinase; enzyme.
XX
OS  Homo sapiens.
XX
PN  US2003049678-A1.
XX
PD  13-MAR-2003.
XX
PF  07-AUG-2001; 2001US-00923830.
XX
PR  05-MAR-1999; 99US-00263904.
XX
PA  (NINEN/) NIENABER V L.
XX  (GREE/) GREER J.
XX  (ABAD/) ABAD-ZAPATERO C.
XX  (NORBE/) NORBECK D W.
XX
PI  Nienaber VL, Greer J, Abad-Zapatero C, Norbeck DW;
XX
DR  WPI; 2003-786861/74.
XX
PT  Identification or designing of ligand to target biomolecule by exposing
XX  target biomolecule crystal to test sample(s), and obtaining X-ray crystal
XX  diffraction pattern to determine, whether ligand/receptor complex is
XX  formed.
XX
PS  Disclosure; Fig 5; 33pp; English.
XX
CC  This invention relates to a novel method for the identification of
CC  ligands to a target biomolecule identified or designed by obtaining a
CC  target biomolecule crystal, exposing the target biomolecule crystal to a
CC  test sample and obtaining X-ray crystal diffraction pattern to determine
CC  whether a ligand-receptor complex is formed. The method is useful for
CC  identifying or designing a ligand to a target biomolecule. The invention
CC  provides direct identification of potential ligands and detailed
CC  information on how the ligand binds and changes in the target. It is
CC  biomolecule. It requires non special labelling of the target. It is
CC  uniquely sensitive to structural changes in the target and the ligand.
CC  The binding event is monitored directly, so that the probability for
CC  false positives is reduced to near zero. The crystallographic data
CC  provide a three dimensional electron density snap-shot of the ligand-
CC  receptor complex showing, which compound binds and how it is bound. The
CC  present sequence is the amino acid sequence of human urokinase which was
CC  used during the exemplification of the invention.
XX
SQ  Sequence 411 AA;

Query Match          100.0%; Score 62; DB 7; Length 411;
Best Local Similarity 100.0%; Pred No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY  1 RRPWCYVOVQ 10
    |||||
    |||||
    |||||
Db  109 RRPWCYVOVQ 118

RESULT 6
AAP93232
ID  AAP93232 standard; peptide; 65 AA.
XX
AC  AAP93232;
XX
DT  25-MAR-2003 (revised)
XX  03-APR-1990 (first entry)
DT

```

XX	Modified scu-PA kringle domain (residues 66-131).
DE	
XX	Plasminogen; activator; t-PA; fibrin; kringle domain;
KM	single-chain urinary plasminogen activator; scu-PA.
KM	
XX	
OS	Homo sapiens.
XX	
PN	W08910401-A.
XX	
PD	02-NOV-1989.
XX	
PE	22-APR-1988; 88US-00184823.
XX	
PR	22-APR-1988; 88US-00184823.
XX	
PA	(COLB) COLLABORATIVE RES INC.
XX	
PI	Mao J, Abercrombi DM;
XX	
DR	WPI; 1989-339965/46.
DR	N-PSDB; AAN92239.
XX	
PT	Modified plasminogen activator - having greater fibrin selectivity and circulating half life.
PT	
XX	
PS	Disclosure; Fig 7; 80pp; English.
XX	
CC	The sequence is encoded by a synthetic oligonucleotide and is a modified scu-PA kringle domain. The modified scu-PA has greater fibrin selectivity and a longer half-life than normal scu-PA. The scu-PA is used in the treatment of pulmonary embolism, thrombosis, myocardial infarct and strokes. (Updated on 25-MAR-2003 to correct PA field.)
CC	
XX	
SQ	Sequence 65 AA;
QY	Query Match 91.9%; Score 57; DB 1; Length 65; Best Local Similarity 100.0%; Pred. No. 0.21; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 RRPWCYGV 9 43 RRPWCYGV 51
RESULT 7	
ABR42617	
ID	ABR42617 standard; protein; 86 AA.
XX	
AC	ABR42617;
XX	
DT	26-AUG-2003 (first entry)
XX	
DE	Human abrogen (hATF-kringle).
XX	
KM	Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor; urokinase plasminogen activator; tumour; metastasis; cytostatic; gene therapy.
KX	
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 43 /label= Asn, Asp
FT	Misc-difference 74 /label= Pro, Leu
FT	
XX	
PN	W02003042354-A2.
XX	
PD	22-MAY-2003.
XX	
PF	04-SEP-2002; 2002WO-US027885.
XX	

PR 04-SEP-2001; 2001US-0316300P.
 XX (AVET) AVENTIS PHARM INC.
 XX
 PI Nesbitt M, Fong TC, Brockstedt D;
 XX WPI; 2003-449566/42.
 DR N-PSDB; ACC58338.
 XX
 PT New abrogen polypeptide, useful for treating an angiogenesis related
 PT diseases e.g. tumor metastasis.
 XX
 PS Disclosure; Page 95; 95pp; English.
 XX
 CC The present sequence is the protein sequence of a novel human abrogen,
 CC designated hATF-kringle, comprising the human urokinase plasminogen
 CC activator kringle domain. Abrogens such as hATF-kringle are potent
 CC inhibitors of endothelial proliferation and angiogenesis. Abrogen
 CC polypeptides are capable of inhibiting or reducing cell proliferation
 CC induced by both basic fibroblast growth factor (bFGF) and vascular
 CC endothelial growth factor in a specific endothelial cell proliferation
 CC assay; angiostatin only inhibits bFGF induced proliferation in this
 CC assay. Vectors that expressed abrogen polypeptides in vivo were shown to
 CC reduce tumor metastasis in 2 lung cancer models. The invention provides
 CC abrogen polypeptides and polynucleotides, and methods of using these to
 CC treat an angiogenesis-related disease or disorder, e.g. tumor metastasis
 CC (claimed)
 CC
 SQ Sequence 86 AA;

Query Match 91.9%; Score 57; DB 7; Length 86;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRPWCYQV 9
 DB 62 RRPWCYQV 70

RESULT 8
 ABR42602
 ID ABR42602 standard; protein; 86 AA.
 XX
 AC ABR42602;
 XX
 DT 26-AUG-2003 (first entry)
 XX
 DE Mouse abrogen (hATF-kringle).
 XX
 KM Mouse; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
 KM urokinase plasminogen activator; tumour; metastasis; cyostatic;
 KM gene therapy.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT Domain 53..59
 FT /note="kringle domain"
 XX
 PN WO2003042354-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 04-SEP-2002; 2002MO-US027885.
 XX
 PR 04-SEP-2001; 2001US-0316300P.
 XX
 PA (AVET) AVENTIS PHARM INC.
 XX
 PI Nesbitt M, Fong TC, Brockstedt D;
 XX WPI; 2003-449566/42.
 DR N-PSDB; ACC58337.

XX
 PT New abrogen polypeptide, useful for treating an angiogenesis related
 PT diseases e.g. tumor metastasis.
 XX
 PS Claim 1; Page 25; 95pp; English.
 XX
 CC The present sequence is the protein sequence of a novel mouse abrogen,
 CC designated mATF-kringle, comprising the human urokinase plasminogen
 CC activator kringle domain. Abrogens such as mATF-kringle are potent
 CC inhibitors of endothelial proliferation and angiogenesis. Abrogen
 CC polypeptides are capable of inhibiting or reducing cell proliferation
 CC induced by both basic fibroblast growth factor (bFGF) and vascular
 CC endothelial growth factor in a specific endothelial cell proliferation
 CC assay; angiostatin only inhibits bFGF induced proliferation in this
 CC assay. Vectors that expressed abrogen polypeptides in vivo were shown to
 CC reduce tumor metastasis in 2 lung cancer models. The invention provides
 CC abrogen polypeptides and polynucleotides, and methods of using these to
 CC treat an angiogenesis-related disease or disorder, e.g. tumor metastasis
 CC (claimed)
 CC
 SQ Sequence 86 AA;

Query Match 91.9%; Score 57; DB 7; Length 86;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRPWCYQV 9
 DB 62 RRPWCYQV 70

RESULT 9
 ABR42599
 ID ABR42599 standard; protein; 86 AA.
 XX
 AC ABR42599;
 XX
 DT 26-AUG-2003 (first entry)
 XX
 DE Human abrogen (hATF-kringle).
 XX
 KM Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
 KM urokinase plasminogen activator; tumour; metastasis; cyostatic;
 KM gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 53..59
 FT /note="kringle domain"
 XX
 PN WO2003042354-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 04-SEP-2002; 2002MO-US027885.
 XX
 PR 04-SEP-2001; 2001US-0316300P.
 XX
 PA (AVET) AVENTIS PHARM INC.
 XX
 PI Nesbitt M, Fong TC, Brockstedt D;
 XX WPI; 2003-449566/42.
 DR N-PSDB; ACC58334.
 XX
 PT New abrogen polypeptide, useful for treating an angiogenesis related
 PT diseases e.g. tumor metastasis.
 XX
 PS Claim 1; Page 24; 95pp; English.
 XX
 CC The present sequence is the protein sequence of a novel human abrogen,
 CC designated hATF-kringle, comprising the human urokinase plasminogen

Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
Db 63 RRPWCYQV 71

RESULT 12

ABR42605 standard; protein; 87 AA.

AC ABR42605;
DT 26-AUG-2003 (first entry)
DE Human abrogen (hATF-kringle).
KW Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW gene therapy.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 54..60
FT /note= "kringle domain"

MO2003042354-A2.

22-MAY-2003.

04-SEP-2002; 2002MO-US027885.

04-SEP-2001; 2001US-0316300P.

(AVET) AVENTIS PHARM INC.

Nesbit M, Fong TC, Brokstedt D;

WPI; 2003-449566/42.

New abrogen polypeptide, useful for treating an angiogenesis related diseases e.g. tumor metastasis.

Claim 37; Page 26; 95pp; English.

The present sequence is the protein sequence of a secreted form of a novel human abrogen, designated hATF-kringle, comprising the human urokinase plasminogen activator kringle domain. The sequence includes an N-terminal alanine residue that results from cleavage of an interleukin-2 signal peptide. Abrogens such as hATF-kringle are potent inhibitors of endothelial proliferation and angiogenesis. Abrogen polypeptides are capable of inhibiting or reducing cell proliferation induced by both basic fibroblast growth factor (bFGF) and vascular endothelial growth factor in a specific endothelial cell proliferation assay; angiostatin only inhibits bFGF induced proliferation in this assay. Vectors that expressed abrogen polypeptides in vivo were shown to reduce tumour metastasis in 2 lung cancer models. The invention provides abrogen polypeptides and polynucleotides, and methods of using these to treat an angiogenesis-related disease or disorder, e.g. tumour metastasis (claimed)

Sequence 87 AA;

Query Match 91.9%; Score 57; DB 7; Length 87;

Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
Db 63 RRPWCYQV 71

RESULT 13

AAE16542 standard; protein; 88 AA.

09-APR-2002 (first entry)

Human urokinase-type plasminogen activator (uPA) kringle.

KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW clotting disorder; uterine contraction disorder; respiratory disease;
KW male impotence; adult respiratory distress syndrome.

OS Homo sapiens.

MO20019752-A2.

27-DEC-2001.

13-JUN-2001; 2001MO-US018976.

20-JUN-2000; 2000US-0212874P.

(TYPE-) UNIV PENNSYLVANIA.

Cines DB, Higazi AA;

WPI; 2002-122240/16.

N-PDB; AAD27075.

Composition for modulating muscle cell and tissue contractility for treating atherosclerosis, asthma, hypertension, glaucoma, impotence, comprising domains from urokinase-type plasminogen activator.

Claim 1; Fig 1A; 117pp; English.

The invention relates to a composition comprising one or more domains of urokinase-type plasminogen activator (uPA). The composition is used to modulate the contractility and angiogenic activity of a mammalian muscle, endothelial cell or tissue. The composition is used for treating stroke, hypertension, thrombotic microangiopathies, surgically induced thrombotic occlusions, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma, diabetic retinopathy, wound healing, clotting disorder, uterine contraction disorder, male impotence, respiratory disease or condition such as asthma, adult respiratory distress syndrome, primary pulmonary hypertension, microvascular thrombotic occlusion, and a disorder associated with chronic intrapulmonary fibrin formation. The present sequence is human urokinase-type plasminogen activator (uPA) kringle

Sequence 88 AA;

Query Match 91.9%; Score 57; DB 5; Length 88;

Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
Db 62 RRPWCYQV 70

RESULT 14

AAW22742 standard; protein; 89 AA.

AAW22742;

XX

DT 12-MAR-1998 (first entry)
 XX
 DE Urokinase residues 43-131.
 XX
 KM Chimeric; human; urinary trypsin; inhibitor; HI-8; cancer; metastasis;
 XX G-domain; urokinase; prevention; leukaemia; lymphoma.
 XX
 OS Homo sapiens.
 XX
 PN MO9725422-A1.
 PD 17-JUL-1997.
 XX
 PF 06-JAN-1997; 97WO-JP000008.
 XX
 PR 08-JAN-1996; 96JP-00001059.
 XX
 PA (NISP) NISSIN FOOD PROD CO LTD.
 PI Kobayashi H, Terao T, Sugino D, Okushima M;
 PS WPI; 1997-372862/34.
 DR
 XX Chimeric protein which inhibits development of metastases in cancer -
 PT contains urinary trypsin inhibitor carboxy-terminal domain linked to
 XX urokinase G-domain.
 PS
 CC Claim 3; Page 72; 97pp; Japanese.
 CC
 CC A novel chimeric protein contains the carboxy-terminal domain of human
 CC urinary trypsin inhibitor (HI-8), which inhibits cancer cell metastasis,
 CC linked to a peptide containing the G-domain of urokinase (AAW22742),
 CC which specifically binds the excess urokinase receptor expressed in
 CC cancer cells. The chimeric protein has the amino-terminal AAW22734, the
 CC carboxy-terminal AAW22735 and a linking sequence selected from AAW22736-
 CC 39 or partial sequences derived from these, specifically AAW38130-63. The
 CC chimeric protein may also have additional amino-terminal sequences
 CC selected from AAW22740 or 9 partial sequences derived from this, and/or
 CC additional carboxy-terminal sequences selected from AAW22743 or 10
 CC partial sequences derived from this. The chimeric protein can be used to
 CC prevent metastasis in, e.g. cancer of the lung, kidney, pancreas,
 CC stomach, colon, rectum, ovary, uterus, brain, skin, muscle, breast or
 CC prostate, and in leukaemia or lymphoma
 CC
 CC
 SQ Sequence 89 AA;
 QY
 DB 1 RRPWCYQV 9
 67 RRPWCYQV 75
 Query Match 91.9%; Score 57; DB 2; Length 89;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 15
 ID AAE16550 standard; protein; 96 AA.
 AC AAE16550;
 DT 09-APR-2002 (first entry)
 DE Human uPA kringle and connecting peptide.
 KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW adult respiratory distress syndrome; male impotence.
 OS Homo sapiens.

XX
 PN MO200197752-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 13-JUN-2001; 2001MO-US018976.
 XX
 PR 20-JUN-2000; 2000US-0212874P.
 XX
 PA (UYPE-) UNIV PENNSYLVANIA.
 PI Cines DB, Higazi AA;
 PS WPI; 2002-122240/16.
 DR N-PDB; AAD27083.
 XX
 PT Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator.
 PS
 CC Claim 25; Fig 11; 117pp; English.
 CC
 CC The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle and
 CC connecting peptide
 CC
 CC
 SQ Sequence 96 AA;
 QY
 DB 1 RRPWCYQV 9
 62 RRPWCYQV 70
 Query Match 91.9%; Score 57; DB 5; Length 96;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Search completed: March 18, 2004, 11:16:23
 Job time : 4.13567 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:13:05 ; Search time 1.15974 Seconds
(without alignments)
445.152 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPWCYQVQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep: *
6: /cgn2_6/prodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	91.9	89	4	US-09-101-272G-62
2	57	91.9	158	2	US-08-797-689-12
3	57	91.9	138	4	US-09-984-186-12
4	57	91.9	194	4	US-09-101-272G-80
5	57	91.9	200	4	US-09-101-272G-73
6	57	91.9	201	4	US-09-101-272G-96
7	57	91.9	208	4	US-09-101-272G-98
8	57	91.9	365	1	US-08-093-741-83
9	57	91.9	365	1	US-08-720-012-83
10	57	91.9	393	2	US-08-560-098A-44
11	57	91.9	393	3	US-08-967-024C-24
12	57	91.9	393	3	US-08-967-024C-25
13	57	91.9	411	1	US-08-087-163-1
14	57	91.9	411	1	US-08-286-748B-18
15	57	91.9	411	1	US-08-153-799-18
16	57	91.9	411	2	US-08-560-098A-48
17	57	91.9	411	2	US-09-181-816-1
18	57	91.9	411	4	US-09-403-736-2
19	57	91.9	430	1	US-07-942-157A-3
20	57	91.9	430	6	5219569-2
21	57	91.9	431	4	US-09-101-272G-1
22	57	91.9	431	6	5188829-1
23	57	91.9	432	2	US-08-560-098A-47
24	57	91.9	432	3	US-08-142-590B-25
25	57	91.9	432	3	US-08-148-910-12
26	57	91.9	432	3	US-08-148-910-12
27	57	91.9	432	3	US-08-148-910-12

28	40	64.5	83	2	US-08-811-949-2	Sequence 2, Appl
29	40	64.5	437	2	US-08-811-949-49	Sequence 49, Appl
30	40	64.5	437	2	US-08-811-949-51	Sequence 51, Appl
31	40	64.5	437	2	US-08-811-949-55	Sequence 55, Appl
32	40	64.5	437	2	US-08-811-949-57	Sequence 57, Appl
33	40	64.5	472	2	US-08-811-949-63	Sequence 63, Appl
34	40	64.5	477	2	US-08-560-098A-51	Sequence 51, Appl
35	40	64.5	527	1	US-07-609-510B-16	Sequence 16, Appl
36	40	64.5	527	2	US-08-811-949-39	Sequence 39, Appl
37	40	64.5	527	5	PCT-US91-01025A-2	Sequence 2, Appl
38	40	64.5	527	6	5185259-8	Sequence 2, Appl
39	40	64.5	527	6	5185259-1	Sequence 2, Appl
40	40	64.5	536	6	5200340-6	Sequence 43, Appl
41	40	64.5	562	2	US-08-811-949-43	Sequence 50, Appl
42	40	64.5	562	2	US-08-560-098A-50	Sequence 38, Appl
43	40	64.5	562	4	US-08-883-795A-38	Sequence 4, Appl
44	40	64.5	562	4	US-09-703-695A-4	Sequence 4, Appl
45	40	64.5	562	6	5185259-3	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-101-272G-62
Sequence 62, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Miasin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.1
SEQ ID NO 62
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: residues 43-131 of the ATP domain of uPA
US-09-101-272G-62

Query Match 91.9%; Score 57; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.0069;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQVQ 9
Db 67 RRPWCYQVQ 75

RESULT 2
US-08-797-689-12
Sequence 12, Application US/08797689
Patent No. 5876969
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Guillon, Jean-Dominique
APPLICANT: Jung, Gerard
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSER: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Colleegeville

STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-12

Query Match 91.9%; Score 57; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 112 RRPWCYQV 120

RESULT 3
US-09-984-186-12
Sequence 12, Application US/09984186
Patent No. 6686179
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSER: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 91.9%; Score 57; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 112 RRPWCYQV 120

RESULT 4
US-09-101-272G-80
Sequence 80, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.1
SEQ ID NO 80
LENGTH: 194
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: ATRFH chimeric protein
US-09-101-272G-80

Query Match 91.9%; Score 57; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
DB 110 RRPWCYQV 118

RESULT 5
US-09-101-272G-73
Sequence 73, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR

FILE REFERENCE: 050979
CURRENT APPLICATION NUMBER: US/09/101.272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin version 3.1
SEQ ID NO 73
LENGTH: 200
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: ATF domain of UPA
US-09-101-272G-73

Query Match 91.9%; Score 57; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 129 RRPWCYQV 137

RESULT 6
US-09-101-272G-96
Sequence 96, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: 050979
CURRENT APPLICATION NUMBER: US/09/101.272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin version 3.1
SEQ ID NO 96
LENGTH: 201
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ATPH1-CL chimeric protein
US-09-101-272G-96

Query Match 91.9%; Score 57; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 110 RRPWCYQV 118

RESULT 7
US-09-101-272G-98
Sequence 98, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: 050979
CURRENT APPLICATION NUMBER: US/09/101.272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin version 3.1
SEQ ID NO 98
LENGTH: 208
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ATPH1-ML chimeric protein
US-09-101-272G-98

Query Match 91.9%; Score 57; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 110 RRPWCYQV 118

RESULT 8
US-08-093-741-83
Sequence 83, Application US/08093741
Patent No. 5681721
GENERAL INFORMATION:
APPLICANT: STEFFENS, GERO J.
APPLICANT: WENDEL, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESSES:
ADDRESSER: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,741
FILING DATE: 20-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-093-741-83

Query Match 91.9%; Score 57; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 63 RRPWCYQV 71

RESULT 9
US-08-720-012-83
Sequence 83, Application US/08720012
Patent No. 5747291

GENERAL INFORMATION:
APPLICANT: STEFFENS, GERD J.
APPLICANT: WENDT, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL-WIELAND, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-720-012-83
Query Match 91.9%; Score 57; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRPWCYQV 9
Db 63 RRPWCYQV 71
RESULT 10
US-08-560-098A-44
Sequence 44, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-44
Query Match 91.9%; Score 57; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRPWCYQV 9
Db 64 RRPWCYQV 72
RESULT 11
US-08-967-024C-24
Sequence 24, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WENDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOCHA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 91.9%; Score 57; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPCTVQV 9
DB 64 RRPCTVQV 72

RESULT 12
US-08-967-024C-25
Sequence 25, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: MENDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOGH, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Evenson, McKee, Edward & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 91.9%; Score 57; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPCTVQV 9
DB 64 RRPCTVQV 72

RESULT 13
US-08-087-163-1
Sequence 1, Application US/08087163
Patent No. 5472692

GENERAL INFORMATION:
APPLICANT: Liu, Jian-Ning
APPLICANT: Gurewich, Victor
TITLE OF INVENTION: PRO-UKINASE MUTANTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50x or 55sx
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,163
FILING DATE: 07/02/93
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Faese, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04353/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
US-08-087-163-1

Query Match 91.9%; Score 57; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPCTVQV 9
DB 109 RRPCTVQV 117

RESULT 14
US-08-286-748B-18
Sequence 18, Application US/08286748B
Patent No. 5759542
GENERAL INFORMATION:
APPLICANT: Victor Gurewich
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50x or 55sx
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,748B

FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Fasse
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-18

Query Match 91.9%; Score 57; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 109 RRPWCYQV 117

RESULT 15

US-08-153-799-18
Sequence 18, Application US/08153799
Patent No. 576683
GENERAL INFORMATION:
APPLICANT: Ballance, David J
APPLICANT: Goodey, Andrew R
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: R Hain Swope, BOC Health Care Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill
STATE: New Jersey
COUNTRY: USA
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400

TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-153-799-18

Query Match 91.9%; Score 57; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 109 RRPWCYQV 117

Search completed: March 18, 2004, 11:21:23
Job time: 2.15974 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:19:25 ; Search time 2.97593 Seconds
(without alignments)
870.166 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPWCYQVQ 10

Scoring table: BL0SUM62
Gap 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	10	13	US-10-076-421-5
2	57	91.9	86	15	US-10-233-675A-1
3	57	91.9	86	15	US-10-233-675A-1
4	57	91.9	86	15	US-10-233-675A-7
5	57	91.9	86	15	US-10-233-675A-22
6	57	91.9	86	15	US-10-233-675A-27
7	57	91.9	87	15	US-10-233-675A-9
8	57	91.9	87	15	US-10-233-675A-10
9	57	91.9	88	9	US-09-880-503-1
10	57	91.9	96	9	US-09-880-503-9
11	57	91.9	135	9	US-09-880-503-4
12	57	91.9	138	9	US-09-984-186-12
13	57	91.9	138	14	US-10-237-667-12
14	57	91.9	138	14	US-10-237-708-12
15	57	91.9	138	14	US-10-237-866-12

16	57	91.9	138	14	US-10-237-871-12	Sequence 12, Appl
17	57	91.9	138	14	US-10-237-624-12	Sequence 12, Appl
18	57	91.9	133	9	US-09-880-503-8	Sequence 8, Appl
19	57	91.9	322	15	US-10-233-675A-20	Sequence 20, Appl
20	57	91.9	322	15	US-10-233-675A-21	Sequence 21, Appl
21	57	91.9	322	15	US-10-233-675A-21	Sequence 22, Appl
22	57	91.9	337	14	US-10-106-698-626	Sequence 626, Ap
23	57	91.9	337	15	US-10-264-049-2927	Sequence 2927, Ap
24	57	91.9	403	9	US-09-880-503-6	Sequence 6, Appl
25	57	91.9	411	15	US-09-880-503-3	Sequence 3, Appl
26	57	91.9	411	15	US-10-407-821-2	Sequence 2, Appl
27	57	91.9	431	12	US-09-264-468B-1	Sequence 1, Appl
28	57	91.9	431	12	US-10-411-037-34	Sequence 34, Appl
29	57	91.9	431	13	US-10-076-421-2	Sequence 2, Appl
30	57	91.9	431	14	US-10-171-311-184	Sequence 184, App
31	57	91.9	431	14	US-10-193-656-4	Sequence 4, Appl
32	57	91.9	431	14	US-10-301-822-161	Sequence 161, App
33	57	91.9	431	14	US-10-247-671-149	Sequence 149, App
34	57	91.9	431	14	US-10-131-985-21	Sequence 21, Appl
35	57	91.9	431	15	US-10-295-027-414	Sequence 414, App
36	57	91.9	445	15	US-10-295-027-1275	Sequence 1275, Ap
37	57	91.9	650	15	US-10-401-077-1	Sequence 266, App
38	57	91.9	672	15	US-10-233-675A-15	Sequence 15, Appl
39	57	91.9	674	15	US-10-233-675A-14	Sequence 14, Appl
40	57	91.9	687	15	US-10-233-675A-17	Sequence 17, Appl
41	57	91.9	688	15	US-10-233-675A-18	Sequence 18, Appl
42	57	91.9	689	15	US-10-233-675A-13	Sequence 13, Appl
43	53	85.5	86	15	US-10-233-675A-3	Sequence 3, Appl
44	53	85.5	86	15	US-10-233-675A-24	Sequence 24, Appl
45	43	69.4	655	14	US-10-172-712-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-10-076-421-5
; Sequence 5, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, MANABU
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-5

Query Match 100.0%; Score 62; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 RRPWCYQVQ 10
Db 1 RRPWCYQVQ 10
RESULT 2
US-10-233-675A-1
; Sequence 1, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: NeebIt, Mark

```
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
US-10-233-675A-1
```

```
Query Match          91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRPWCYQV 9
    |||||
Db 62 RRPWCYQV 70
```

```
RESULT 3
US-10-233-675A-5
; Sequence 5, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
US-10-233-675A-5
```

```
Query Match          91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRPWCYQV 9
    |||||
Db 62 RRPWCYQV 70
```

```
RESULT 4
US-10-233-675A-7
; Sequence 7, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
```

```
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
US-10-233-675A-7
```

```
Query Match          91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRPWCYQV 9
    |||||
Db 62 RRPWCYQV 70
```

```
RESULT 5
US-10-233-675A-22
; Sequence 22, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human urokinase plasminogen activator
US-10-233-675A-22
```

```
Query Match          91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRPWCYQV 9
    |||||
Db 62 RRPWCYQV 70
```

```
RESULT 6
US-10-233-675A-27
; Sequence 27, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
```

```
SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human urokinase plasminogen activator
; NAME/KEY: MISC FEATURE
; LOCATION: (43)-(43)
; OTHER INFORMATION: Polymorphism - Xaa = Asn or Asp
; NAME/KEY: MISC FEATURE
; LOCATION: (74)-(74)
; OTHER INFORMATION: Polymorphism - Xaa = Pro or Leu
US-10-233-675A-27
```

```
Query Match          91.9%; Score 57; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RRPWCYQV 9
        |||||
Db       62 RRPWCYQV 70
```

```
RESULT 7
US-10-233-675A-9
; Sequence 9, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived fusion protein
US-10-233-675A-9
```

```
Query Match          91.9%; Score 57; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RRPWCYQV 9
        |||||
Db       63 RRPWCYQV 71
```

```
RESULT 8
US-10-233-675A-10
; Sequence 10, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
```

```
PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived fusion protein
US-10-233-675A-10
```

```
Query Match          91.9%; Score 57; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RRPWCYQV 9
        |||||
Db       63 RRPWCYQV 71
```

```
RESULT 9
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1
```

```
Query Match          91.9%; Score 57; DB 9; Length 88;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 RRPWCYQV 9
        |||||
Db       62 RRPWCYQV 70
```

```
RESULT 10
US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-9
```

Query Match 91.9%; Score 57; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRPWCYQV 9
|||
Db 62 RRPWCYQV 70

RESULT 11
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Rooof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match 91.9%; Score 57; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRPWCYQV 9
|||
Db 109 RRPWCYQV 117

RESULT 12
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689

FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: S792006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 91.9%; Score 57; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRPWCYQV 9
|||
Db 112 RRPWCYQV 120

RESULT 13
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US2003022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,667
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: S792006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12

Query Match
Best Local Similarity 91.9%; Score 57; DB 14; Length 138;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
112 RRPWCYQV 120

Db 112 RRPWCYQV 120

RESULT 14
US-10-237-708-12
Sequence 12, Application US/10237708
Publication No. US20030036170A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: S792006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match
Best Local Similarity 91.9%; Score 57; DB 14; Length 138;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
112 RRPWCYQV 120

Db 112 RRPWCYQV 120

RESULT 15
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: S792006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match
Best Local Similarity 91.9%; Score 57; DB 14; Length 138;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Mon Mar 22 09:48:10 2004

us-10-076-421-5.rapb

Page 6

OY 1 RRPWCYVQV 9
| | | | |
Db 112 RRPWCYVQV 120

Search completed: March 18, 2004, 11:36:07
Job time : 3.22593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 11:10:10 ; Search time 1.02845 Seconds
(without alignments)
935.309 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPWCYVQVQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	91.9	431	1 UKHU	u-plasminogen acti
2	57	91.9	433	1 UKBAY	u-plasminogen acti
3	57	91.9	442	1 UKPG	u-plasminogen acti
4	56	90.3	432	1 S18932	u-plasminogen acti
5	56	90.3	433	1 JN0560	u-plasminogen acti
6	53	85.5	433	1 UKMS	u-plasminogen acti
7	43	69.4	608	2 G82137	pyca protein VC194
8	43	69.4	655	1 A46688	hepatocyte growth
9	41	66.1	558	2 JCS878	plasma hyaluronan
10	40	64.5	291	2 J38098	conserved hypochet
11	40	64.5	304	2 B87639	u-plasminogen acti
12	40	64.5	394	2 J50600	u-plasminogen acti
13	40	64.5	431	2 J50589	u-plasminogen acti
14	40	64.5	434	1 A35005	u-plasminogen acti
15	40	64.5	477	1 A34369	u-plasminogen acti
16	40	64.5	477	2 J50597	u-plasminogen acti
17	40	64.5	477	2 J50598	u-plasminogen acti
18	40	64.5	559	1 A35029	u-plasminogen acti
19	40	64.5	559	1 A29941	u-plasminogen acti
20	40	64.5	560	1 JC4795	plasma hyaluronan
21	40	64.5	562	1 UKHU	u-plasminogen acti
22	40	64.5	1113	1 A47106	myosin heavy chain
23	39	62.9	259	2 PN0687	cyclin - antiracno
24	39	62.9	411	2 I51285	hepatocyte growth
25	39	62.9	603	2 S28941	coagulation factor
26	39	62.9	615	1 KFHU12	coagulation factor
27	39	62.9	710	1 I51283	hepatocyte growth
28	39	62.9	711	1 A47136	macrophage-stimula
29	39	62.9	716	1 A40332	macrophage-stimula

30	39	62.9	716	1 JCS061	macrophage-stimula
31	39	62.9	728	1 JH0579	hepatocyte growth
32	39	62.9	728	1 A35644	hepatocyte growth
33	39	62.9	728	1 A60185	hepatocyte growth
34	39	62.9	806	2 T18840	hypothetical prote
35	39	62.9	946	1 A47239	hypothetical prote
36	39	62.9	1420	2 A32869	hypothetical prote
37	39	62.9	4548	1 A30657	hypothetical prote
38	38	61.3	89	2 A60140	plasma (EC 3.4.21
39	38	61.3	323	2 T25094	plasma (EC 3.4.21
40	38	61.3	593	2 S45281	hypothetical prote
41	37	59.7	157	2 T02034	coagulation factor
42	37	59.7	120	2 T46524	early light-induce
43	37	59.7	252	2 E75491	probable disulfide
44	37	59.7	252	2 E75491	fine protein - Dei
45	37	59.7	656	2 S69554	probable fadB17 pr
					hypothetical prote

ALIGNMENTS

RESULT 1
UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N/Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N/Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen acti
in form
C/Species: Homo sapiens (man)
C/Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C/Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A37565
R/Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasl, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A/Title: The human urokinase-plasminogen activator gene and its promoter.
A/Reference number: A00931; MUID:85215647; PMID:2987867
A/Accession: A00931
A/Molecule type: DNA
A/Residues: 1-431 <RC>
A/Cross-references: GB:K03427; NID:937601; PIDN:CAA26268.1; PID:G1834524
A/Note: the authors translated the codon ATG for residue 214 as Ile
R/Nagamine, Y.; Pearson, D.; Gratian, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A/Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine
A/Reference number: I52209; MUID:86050639; PMID:3933505
A/Accession: I52209
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 145-161 <NAG1>
A/Cross-references: NID:9340174; PIDN:AAA61257.1; PID:9340175
R/Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasaka, N.; Aritama, H.; Nishida, M.; Suyama, J
Gene 36, 183-188, 1985
A/Title: Molecular cloning of cDNA coding for human preprourokinase.
A/Reference number: J70102; MUID:86056954; PMID:2415429
A/Accession: J70102
A/Molecule type: mRNA
A/Residues: 1-213, 'I', 215-431 <NAG2>
A/Cross-references: GB:K03226; NID:9340155; PIDN:AA971138.1; PID:9340158; GB:DD0244; NID
R/Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Biasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A/Title: Identification and primary sequence of an unspliced human urokinase poly(A)+ RN
A/Reference number: A37561; MUID:84272706; PMID:6589620
A/Accession: A37561
A/Molecule type: mRNA
A/Residues: 66-431 <VER>
A/Cross-references: GB:DD0244; NID:9220138
R/Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen,
DNA 4, 139-146, 1985
A/Title: Molecular cloning, sequencing, and expression in Escherichia coli of human prepr
A/Reference number: I38102; MUID:85203359; PMID:3888571
A/Accession: I38102
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <VAC>
A/Cross-references: EMBL:X02760; NID:935297; PIDN:CAA26535.1; PID:935298

A;Experimental source: kidney cell line LLC-PK1
R;Nagamine, Y.
Submitted to the Protein Sequence Database, December 1986
A;Reference number: A37566
A;Contents: annotation; correction to residue 241
C;Genetics:
A;Intons: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-168/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;33-64/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>
F;190-430/Domain: trypsin homology <TRY>
F;152/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;119-310,220-236,228-299,334-393,356-372,383-411/Disulfide bonds: #status predicted
F;235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 91.9%; Score 57; DB 1; Length 442;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 131 RRPWCYQV 139

RESULT 4
S18932
U-plasminogen activator (EC 3.4.21.73) precursor - rat
N;Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C;Species: Rattus norvegicus (Norway rat)
C;Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text_change 18-Jun-1999
C;Accession: S24604; I60186; I53472; S18932
R;Rabban, S.A.
Submitted to the EMBL Data Library, April 1992
A;Reference number: S24604
A;Accession: S24604
A;Molecule type: mRNA
A;Residues: 1-15, 'H', 'I', '23', 'G', '25-331', 'N', '333-432' <RAB>
A;Cross-references: EMBL:X65651; NID:957456; PIDN:CAA4601.1; PID:957457
A;Experimental source: tissue kidney
R;Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A;Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
A;Reference number: I60186; MUID:92233409; PMID:1568219
A;Accession: I60186
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-432 <RES>
A;Cross-references: EMBL:X63434; NID:957465; PIDN:CAA5028.1; PID:957466
A;Experimental source: strain Fischer 344; tissue mammary
R;Ragno, P.; Casasno, S.; Degen, J.; Kessler, C.; Blas, F.; Rosel, G.
FEBS Lett. 306, 193-198, 1992
A;Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
A;Reference number: I53472; MUID:92333549; PMID:1321734
A;Accession: I53472
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 31-62 <RE2>
A;Cross-references: EMBL:X66907; NID:9396200; PIDN:CAA47356.1; PID:938279
C;Genetics:
A;Gene: uPA
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;119-332/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>
F;119-420/Domain: trypsin homology <TRY>
F;168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted

F;225,276,377/Active site: His, Asp, Ser #status predicted

Query Match 90.3%; Score 56; DB 1; Length 432;
Best Local Similarity 88.9%; Pred. No. 0.067;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 129 RRPWCYQV 137

RESULT 5
JN0560
U-plasminogen activator (EC 3.4.21.73) precursor - bovine
N;Alternate names: uPA
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 16-Jul-1999
C;Accession: JN0560
R;Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A;Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induc
A;Reference number: JN0560; MUID:93216119; PMID:8385052
A;Accession: JN0560
A;Molecule type: mRNA
A;Residues: 1-433 <KRA>
A;Cross-references: GB:I03546; NID:9163800; PIDN:AAA51419.1; PID:9163801
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-179/Product: plasminogen activator chain A #status predicted <MA1>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KRG>
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>
F;181-421/Domain: trypsin homology <TRY>
F;170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 90.3%; Score 56; DB 1; Length 433;
Best Local Similarity 88.9%; Pred. No. 0.067;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
Db 131 RRPWCYQV 139

RESULT 6
U005
U-plasminogen activator (EC 3.4.21.73) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1987 #sequence revision 30-Sep-1987 #text_change 18-Jun-1999
C;Accession: A29420; A24615
R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
Biochemistry 26, 8270-8279, 1987
A;Title: The murine urokinase-type plasminogen activator gene.
A;Reference number: A29420; MUID:86163489; PMID:2831940
A;Accession: A29420
A;Molecule type: DNA
A;Residues: 1-433 <DEG>
A;Cross-references: GB:M17922; NID:9202296; PIDN:AAA40539.1; PID:9202297
R;Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher,
Eur. J. Biochem. 148, 225-232, 1985
A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase-
A;Reference number: A24615; MUID:85179474; PMID:2985383
A;Accession: A24615
A;Molecule type: mRNA
A;Residues: 1-433 <BEL>
A;Cross-references: GB:X02389; NID:955127; PIDN:CAA26231.1; PID:955128
C;Genetics:
A;Intons: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F.1-20/Domain: signal sequence #status predicted <SIG>
F.21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F.32-63/Domain: EGF homology <EGF>
F.71-152/Domain: kringle homology <KR>
F.180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>
F.180-421/Domain: trypsin homology <TRY>
F.169-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F.226-277,378/Active site: His, Asp, Ser #status predicted

Query Match 85.5%; Score 53; DB 1; Length 433;
Best Local Similarity 77.8%; Pred. No. 0.21;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
:|||||:
Db 130 KRPWCYQV 138

RESULT 7
G82137
pvcA protein VC1949 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82137
R:Heidelberg, V.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Drygiol, I.; Sellers, H.
I. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nucleotide 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82137
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-608 <HEI>
A:Cross-references: GB:AE004270; GB:AE003852; NID:g9656479; PIDN:AA95097.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
A:Genetics:
A:Gene: VC1949
A:Map position: 1

Query Match 69.4%; Score 43; DB 2; Length 608;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPPWCYQV 9
:|||||:
Db 306 KPPWCYQV 313

RESULT 8
A46688
hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
C:Species: Homo sapiens (man)
C>Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C:Accession: A46688
R:Miyaizawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
d coagulation factor XII
A:Reference number: A46688; MUID:93252878; PMID:7683665
A:Accession: A46688
A:Molecule type: mRNA
A:Residues: 1-655 <MIY>
A:Cross-references: DDBJ:D14012; NID:g219680; PIDN:BA03113.1; PID:g219681
A:Experimental source: liver (mRNA); serum (protein)
A:Note: sequence extracted from NCBI backbone (NCBI:131227, NCBI:131228)
A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c
C:Genetics:
A:Gene: GDB:HGFA; HGFA; HGFA
A:Cross-references: GDB:9954514
A:Map position: 4p16-4p16
C:Function:
A:Description: activates hepatocyte growth factor by specific proteolytic cleavage

A:Pathway: tissue repair and regeneration
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F.1-34/Domain: signal sequence #status predicted <SIG>
F.108-148/Domain: fibronectin type II repeat homology <IF2>
F.164-157/Domain: EGF homology <EG1>
F.202-237/Domain: fibronectin type I repeat homology <IF1>
F.245-278/Domain: EGF homology <EG2>
F.286-367/Domain: kringle homology <KR>
F.373-407/Product: hepatocyte growth factor activator light chain #status experimental <L
F.408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <H
F.406-641/Domain: trypsin homology <TRY>
F.40-48,290,468,492,546/Binding site: carbohydrate (Aan) (covalent) #status predicted
F.164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-362
F.447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 69.4%; Score 43; DB 1; Length 655;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPPWCYV 7
:|||||:
Db 346 RPPWCYV 351

RESULT 9
JC5878
Plasma hyaluronan-binding protein precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C:Accession: JC5878
R:Hashimoto, K.; Tohe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, N.J
Biol. Pharm. Bull. 20, 1127-1130, 1997
A:Title: Cloning of the cDNA for a mouse homologue of human PHB: A novel hyaluronan-bin
A:Reference number: JC5878; MUID:98065239; PMID:9401717
A:Accession: JC5878
A:Molecule type: mRNA
A:Residues: 1-558 <HAS>
A:Comment: This protein acts as serine protease.
C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsi
F.24-311/Domain: signal sequence #status predicted <SIG>
F.24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>
F.75-106/Domain: EGF homology <EG1>
F.113-145/Domain: EGF homology <EG2>
F.152-185/Domain: EGF homology <EG3>
F.192-274/Domain: kringle homology <KR>
F.312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MAT
F.312-548/Domain: trypsin homology <TRY>

Query Match 66.1%; Score 41; DB 2; Length 558;
Best Local Similarity 62.5%; Pred. No. 25;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPPWCYQV 9
:|||||:
Db 252 KPPWCYQV 259

RESULT 10
I38098
t-plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: tissue plasminogen activator
C:Species: Homo sapiens (man)
C>Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res. 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PvAT) cDNA obtained from human endo.
A:Reference number: I38098; MUID:90192128; PMID:1969145
A:Accession: I38098
A:Status: translated from GB/EMBL/DDBT
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>

A/Cross-references: EMBL:X13097; NID:G35282; PIDN:CAA31489.1; PID:G35283
 C/Comment: For the main splice form, see PIR:UKHTT. This form probably does not have pro
 C/Genetics:
 A/Gene: GDB:PLAT
 A/Cross-references: GDB:119496; OMIM:173370
 A/Map position: BP12-8p12
 A/Intons: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
 C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C/Keywords: alternative splicing; fibrinolytic; glycoprotein; kringle
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-32/Domain: propeptide #status predicted <PRO>
 F:33-251/Product: t-plasminogen activator; inactive endothelial splice form #status pred
 F:41-78/Domain: fibronectin type I repeat homology <IFA>
 F:86-119/Domain: EGF homology <EGF>
 F:127-208/Domain: kringle homology <KR1>
 F:215-291/Domain: kringle homology #status atypical <KR2>
 F:41-71.69-78.86-97.91-108.110-115.127-208.148-190.179-203/Disulfide bonds: #status pred

Query Match 64.5%; Score 40; DB 2; Length 291;
 Best Local Similarity 83.3%; Pred. No. 21;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPKCYV 7
 :|||||
 Db 187 KPMCYV 192

RESULT 11
 conserved hypothetical protein CC3148 [imported] - Caulobacter crescentus
 C/Species: Caulobacter crescentus
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C/Accession: B87639
 R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Neilson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.D.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor
 n, U.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A/Title: Complete Genome Sequence of Caulobacter crescentus.
 A/Reference number: A87249; MUID:21173698; PMID:11259647
 A/Accession: B87639
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-304 <STO>
 A/Cross-references: GB:AE005673; NID:G13424816; PIDN:AAK25110.1; GSPDB:GN00148
 C/Genetics:
 A/Gene: CC3148

Query Match 64.5%; Score 40; DB 2; Length 304;
 Best Local Similarity 75.0%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPKCYV 10
 :|||||
 Db 25 PMLCVQR 32

RESULT 12
 JS0600
 t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
 N/Alternate names: tissue plasminogen activator
 C/Species: Desmodus rotundus (common vampire bat)
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C/Accession: JS0600
 R/Kraetzschmar, J.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon, A.; Dor
 Gene 105, 229-237, 1991
 A/Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A/Reference number: JS0597; MUID:92039036; PMID:1937019
 A/Accession: JS0600
 A/Molecule type: mRNA
 A/Residues: 1-394 <KRA>
 A/Cross-references: GB:M63990; NID:G166078; PIDN:AAA31595.1; PID:G166079
 A/Note: the authors translated the codon ATC for residue 75 as Thr
 C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom

C/Keywords: fibrinolysis; glycoprotein; hydrolyase; kringle; serine proteinase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-36/Domain: propeptide #status predicted <PRO>
 F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
 F:45-126/Domain: kringle homology <KR2>
 F:143-388/Domain: trypsin homology <TRY>
 F:45-126,66-108,97-121,131-262,174-190,192-251,276-351,308-324,341-369/Disulfide bonds:
 F:142-153/Cleavage site: His-Ser (plasmin) #status predicted
 F:189,228,345/Active site: His, Asp, Ser #status predicted
 F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 64.5%; Score 40; DB 2; Length 394;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPKCYV 7
 :|||||
 Db 105 KPMCYV 110

RESULT 13
 JS0599
 t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
 N/Alternate names: tissue plasminogen activator
 C/Species: Desmodus rotundus (common vampire bat)
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C/Accession: JS0599
 R/Kraetzschmar, J.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon, A.; Don
 Gene 105, 229-237, 1991
 A/Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A/Reference number: JS0597; MUID:92039036; PMID:1937019
 A/Accession: JS0599
 A/Molecule type: mRNA
 A/Residues: 1-431 <KRA>
 A/Cross-references: GB:M63989; NID:G166076; PIDN:AAA31594.1; PID:G166077
 C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C/Keywords: fibrinolysis; glycoprotein; hydrolyase; kringle; serine proteinase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-36/Domain: propeptide #status predicted <PRO>
 F:37-431/Product: plasminogen activator beta #status predicted <PLA>
 F:41-74/Domain: EGF homology <EGF>
 F:82-163/Domain: kringle homology <KR2>
 F:160-425/Domain: trypsin homology <TRY>
 F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bonds:
 F:119,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:179-180/Cleavage site: His-Ser (plasmin) #status predicted
 F:226,275,382/Active site: His, Asp, Ser #status predicted
 F:345-361,378-406/Disulfide bonds: #status predicted

Query Match 64.5%; Score 40; DB 2; Length 431;
 Best Local Similarity 83.3%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPKCYV 7
 :|||||
 Db 142 KPMCYV 147

RESULT 14
 A35005
 u-plasminogen activator (EC 3.4.21.73) precursor - chicken
 N/Alternate names: uPA
 C/Species: Gallus gallus (chicken)
 C/Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
 C/Accession: A35005
 J/Leille, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
 J. Biol. Chem. 265, 1339-1344, 1990
 A/Title: The chicken urokinase-type plasminogen activator gene.
 A/Reference number: A35005; MUID:90110185; PMID:2255632
 A/Accession: A35005
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-434 <LES>

A;Cross-references: GB:J05187, NID:g212858, PIDN:AAA9131.1, PID:g212859
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F;40-71/Domain: EGF homology <EGF>
 F;79-158/Domain: Kringle homology <KR>
 F;173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F;173-416/Domain: trypsin homology <TRY>
 F;162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted
 F;217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 64.5%; Score 40; DB 1; Length 434;
 Best Local Similarity 71.4%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RPMCYVQ 8
 :|||||:
 Db 136 RPMCYTK 142

RESULT 15

A34369
 t-plasminogen activator (BC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
 C;Species: Megaderma lyra
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: A34369
 R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
 J. Biol. Chem. 264, 17947-17952, 1989
 A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
 A;Reference number: A34369; MUID:90036867; PMID:2509450
 A;Accession: A34369
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-477 <GAR>
 A;Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <PRO>
 F;22-36/Domain: propeptide #status predicted <PLA>
 F;37-477/Product: plasminogen activator #status predicted <LFA>
 F;42-79/Domain: fibronectin type I repeat homology <LFA>
 F;87-120/Domain: EGF homology <EGF>
 F;128-209/Domain: kringle homology <KR>
 F;226-477/Domain: trypsin homology <TRY>
 F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
 F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 64.5%; Score 40; DB 1; Length 477;
 Best Local Similarity 83.3%; Pred. No. 32;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPMCYV 7
 :|||||:
 Db 188 RPMCYV 193

Search completed: March 18, 2004, 11:20:19
 Job time : 2.02845 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 10:54:45 ; Search time 0.612691 Seconds
(without alignments)
849.859 Million cell updates/sec

Title: US-10-076-421-5

Perfect score: 62

Sequence: 1 RRPACTYQVQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	57	91.9	431 1	UROC_HUMAN
2	57	91.9	433 1	UROC_PAPCY
3	57	91.9	442 1	UROC_PIG
4	56	90.3	432 1	UROC_RAT
5	56	90.3	433 1	UROC_BOVIN
6	53	85.5	433 1	UROC_MOUSE
7	43	69.4	653 1	HGFA_MOUSE
8	43	69.4	655 1	HGFA_HUMAN
9	40	64.5	394 1	URTB_DESRO
10	40	64.5	431 1	URTB_DESRO
11	40	64.5	434 1	UROC_CHICK
12	40	64.5	477 1	URTB_DESRO
13	40	64.5	477 1	URTB_DESRO
14	40	64.5	559 1	TPA_MOUSE
15	40	64.5	559 1	TPA_RAT
16	40	64.5	562 1	TPA_HUMAN
17	40	64.5	566 1	TPA_BOVIN
18	40	64.5	1109 1	MYSD_DICDI
19	39	62.9	245 1	CGI_COLCI
20	39	62.9	452 1	KRMI_XENLA
21	39	62.9	461 1	KRMI_MOUSE
22	39	62.9	462 1	KRMI_HUMAN
23	39	62.9	603 1	FA12_CAVPO
24	39	62.9	615 1	FA12_HUMAN
25	39	62.9	711 1	HGFL_HUMAN
26	39	62.9	716 1	HGFL_MOUSE
27	39	62.9	728 1	HGF_HUMAN
28	39	62.9	728 1	HGF_MOUSE
29	39	62.9	728 1	HGF_RAT
30	39	62.9	1420 1	APOA_MACMU
31	39	62.9	4548 1	APOA_HUMAN
32	38	61.3	473 1	KRMI_MOUSE
33	38	61.3	473 1	KRMI_RAT

34	38	61.3	475 1	KRMI_HUMAN	Q6em8 homo sapien
35	38	61.3	593 1	FA12_BOVIN	P8140 bos taurus
36	37.5	60.5	862 1	TASY_TAXBA	Q93ya3 taxus bacc
37	37.5	60.5	862 1	TASY_TAXBR	Q41594 taxus brevi
38	37.5	60.5	862 1	TASY_TAXCH	Q9ftc37 taxus chine
39	37	59.7	416 1	ORF5_STRGR	Q54238 streptomyc
40	37	59.7	418 1	Y516_STRCO	Q9k414 streptomyc
41	37	59.7	512 1	BETC_RHIME	Q69787 rhizobium m
42	37	59.7	656 1	PUP6_YEAST	Q04373 saccharomyc
43	36	58.1	333 1	PLMN_CANPA	P80009 canis famli
44	36	58.1	374 1	SPOB_HUMAN	Q43791 homo sapien
45	36	58.1	758 1	CC27_YEAST	P38042 saccharomyc

ALIGNMENTS

RESULT 1
UROC_HUMAN STANDARD; PRT; 431 AA.
ID UROC_HUMAN AC P00749; Q15844; Q16618; Q969W6;
DT 21-JUN-1986 (Ref. 01, Created)
DT 20-MAR-1987 (Ref. 04, Last sequence update)
DT 10-OCT-2003 (Ref. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter."; Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RT Steffens G.J., Heyneker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli."; Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=6056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
RT "Molecular cloning of cDNA coding for human prepro-urokinase."; Gene 36:183-186(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Cravador A., Lortien R., Brockly F., Colau B., Chuchana P.,
RA van Elsen A., Herzog A., Bollen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA."; DNA 4:135-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Peingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W.,
RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [17]
RN SEQUENCE OF 66-431 FROM N.A.
RX MEDLINE=84272706; PubMed=6589620;
RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.,
RT "Identification and primary sequence of an unspliced human urokinase
RT poly(A)+ RNA.";
RT Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
RL [8]
RN SEQUENCE OF 21-177.
RX MEDLINE=83055084; PubMed=6754569;
RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
RA Flohe L.,
RT "The primary structure of high molecular mass urokinase from human
RT urine. The complete amino acid sequence of the A chain.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
RN [9]
RN SEQUENCE OF 156-176 AND 179-224.
RX MEDLINE=83003608; PubMed=6749491;
RA Schaller R.O., Nick H., Rickli E.E., Gilleesen D., Lergier W.,
RA Studer R.O.,
RT "Human low-molecular-weight urinary urokinase. Partial
RT characterization and preliminary sequence data of the two polypeptide
RT chains.";
RL Eur. J. Biochem. 125:251-257(1982).
RN [10]
RN SEQUENCE OF 158-410.
RX MEDLINE=83055099; PubMed=6754572;
RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.,
RT "The complete amino acid sequence of low molecular mass urokinase
RT from human urine.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
RN [11]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=96000858; PubMed=8591045;
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RA Dobson C.M., Stuart D.I., Jones E.Y.,
RT "The crystal structure of the catalytic domain of human
RT urokinase-type plasminogen activator.";
RT Structure 3:681-691(1995).
RN [12]
RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RX MEDLINE=20266327; PubMed=10805774;
RA Speil S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
RA Bode W., Magdolen V., Huber R., Moroder L.,
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
RT selective inhibitors of human urokinase.";
RT Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
RN [13]
RN STRUCTURE BY NMR.
RX MEDLINE=89127526; PubMed=2536903;
RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.,
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
RT dimensional NMR.";
RL Nature 337:579-582(1989).
RN [14]
RN STRUCTURE BY NMR OF 67-155.
RX MEDLINE=93003110; PubMed=1327118;
RL Li X., Smith R.A.G., Dobson C.M.,

RT "Sequential 1H NMR assignments and secondary structure of the kringle
RT domain from urokinase.";
RL Biochemistry 31:9562-9571(1992).
RN [15]
RN STRUCTURE BY NMR OF 67-155.
RX MEDLINE=9419701; PubMed=8107091;
RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.,
RT "Solution structure of the kringle domain from urokinase-type
RT plasminogen activator.";
RL J. Mol. Biol. 235:1548-1559(1994).
RN [16]
RN VARIANT LEU-141.
RX MEDLINE=96186279; PubMed=8652631;
RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
RA Sawasaki Y., Hanada K.,
RT "Characterization of single chain urokinase-type plasminogen
RT activator with a novel amino-acid substitution in the kringle
RT structure.";
RL Biochim. Biophys. Acta 1293:83-89(1996).
RN [17]
RN VARIANT LEU-141.
RX MEDLINE=97218551; PubMed=9065988;
RA Comme B., Berczy M., Belin D.,
RT "Detection of polymorphisms in the human urokinase-type plasminogen
RT activator gene.";
RL Thromb. Haemost. 77:434-435(1997).
RN [18]
RN ERRATUM.
RA Comme B., Berczy M., Belin D.,
RL Thromb. Haemost. 78:973-973(1997).
RN [19]
RN VARIANT LEU-141.
RX MEDLINE=97337920; PubMed=9194591;
RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
RA Creutzburg S., Graeff H., Magdolen V.,
RT "Mutational analysis of the genes encoding urokinase-type plasminogen
RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
RL Electrophoresis 18:686-689(1997).
CC -1- FUNCTION: Potent plasminogen activator and is clinically used for
CC therapy of thrombolytic disorders.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 155 in the low
CC molecular mass form to yield a short A1 chain.
CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to licences@isb-sib.ch).
CC -----
DR EMBL; X02419; CA26268.1; -;
DR EMBL; M15476; AAA61253.1; -;
DR EMBL; D00244; BAA00175.1; -;
DR EMBL; D11443; BAA01919.1; -;
DR EMBL; X02760; CAA26535.1; -;
DR EMBL; AF377330; AAK53822.1; -;
DR EMBL; BC03266; AAC97138.1; -;
DR EMBL; K03286; AAA61252.1; -;
DR EMBL; A21571; CAA01559.1; -;
DR EMBL; A18397; CAA01390.1; -;
DR PIR; A00931; UKHU.
DR PDB; 1KDU; 31-OCT-93.

Query Match 91.9%; Score 57; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
 DB 129 RRPWCYQV 137

RESULT 2
 UROK_PAPCY STANDARD; PRT; 433 AA.
 AC P16227;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 GN (U-plasminogen activator).
 OS Papio cynocephalus (Yellow baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 NCBI_TaxID=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thoracic aorta;
 RX MEDLINE=90287734; PubMed=2113276;
 RA Au Y.P.T., Wang T.W., Clowes A.W.;
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
 RT plasminogen activator";
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 155 in the low
 CC molecular mass form to yield a short A1 chain (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

EMBL: X51935, CA36200.1, -.
 DR PIR; S14687; URGAY.
 DR HSSP; P00749; ILMW.
 DR MEROPS; S01.231; -.
 DR InterPro; IPRO09003; Cys Ser trypsin.
 DR InterPro; IPRO06209; EGF-like.
 DR InterPro; IPRO06210; IEGF.
 DR InterPro; IPRO00001; Kringle.
 DR InterPro; IPRO08293; Pept_S1A_uPA.
 DR InterPro; IPRO01254; Peptidase_S1.
 DR InterPro; IPRO01314; Peptidase_S1A.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PIRSF; PIRSF001144; URK_plasm act; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.

DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR PLASMINOGEN activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
 FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
 FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 26 62 EGF-LIKE.
 FT DOMAIN 69 150 KRINGLE.
 FT DOMAIN 151 177 CONNECTING PEPTIDE.
 FT DISULFID 178 433 SERINE PROTEASE.
 FT DISULFID 30 38 BY SIMILARITY.
 FT DISULFID 32 50 BY SIMILARITY.
 FT DISULFID 52 61 BY SIMILARITY.
 FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 208 224 BY SIMILARITY.
 FT DISULFID 216 287 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT ACT_SITE 374 402 BY SIMILARITY.
 FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
 FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
 FT CATHOD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
 SQ SEQUENCE 433 AA; 48595 MW; 816D22DFDCC8792 CRC64;

Query Match 91.9%; Score 57; DB 1; Length 433;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPWCYQV 9
 DB 128 RRPWCYQV 136

RESULT 3
 UROK_PIG STANDARD; PRT; 442 AA.
 AC P04185;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 GN (U-plasminogen activator).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=85087954; PubMed=6096832;
 RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
 RT "cDNA and gene nucleotide sequence of porcine plasminogen activator";
 RL Nucleic Acids Res. 12:9525-9541(1984).
 GN [2]
 RP REVISION TO 241.
 RA Nagamine Y.;
 RL Submitted (DEC-1986) to the PIR data bank.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).

EMBL; X01648; CAA25806.1; --
 EMBL; X02724; CAA26511.1; --
 PIR; A00932; UKPG.
 HSSP; P00749; 1KDU.
 MEROPS; S01.231; --
 InterPro; IPR009003; Cys Ser trypsin.
 InterPro; IPR006209; EGF-like.
 InterPro; IPR000001; Kringle.
 InterPro; IPR008293; Pept_SIA_uPA.
 InterPro; IPR01254; Peptidase_S1.
 InterPro; IPR01314; Peptidase_S1A.
 Pfam; PF00051; kringle; 1.
 Pfam; PF00089; trypsin; 1.
 PIRSF; PIRSF01144; UKR_plasm_act; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 ProDom; PD000395; Kringle; 1.
 SMART; SMO0130; KR; 1.
 SMART; SMO0020; Tryp_Spc; 1.
 PROSITE; PS00022; EGF_1; 1.
 PROSITE; PS01186; EGF_2; FALSE_NEG.
 PROSITE; PS50026; EGF_3; 1.
 PROSITE; PS00021; KRINGLE_1; 1.
 PROSITE; PS50070; KRINGLE_2; 1.
 PROSITE; PS50240; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 KMW; Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 Kringle; EGF-like domain; Zymogen; Signal.
 SIGNAL; 1; 20
 CHAIN; 21; 442
 CHAIN; 190; 188
 CHAIN; 190; 442
 DOMAIN; 29; 65
 DOMAIN; 72; 153
 DOMAIN; 154; 189
 DOMAIN; 190; 442
 CAROHD; 152; 152
 DISULFID; 33; 41
 DISULFID; 35; 53
 DISULFID; 55; 64
 DISULFID; 179; 310
 DISULFID; 220; 236
 DISULFID; 228; 299
 DISULFID; 324; 393
 DISULFID; 356; 372
 DISULFID; 363; 411
 DISULFID; 411; 235
 ACT_SITE; 286; 286
 ACT_SITE; 286; 286
 ACT_SITE; 287; 387
 CONFLICT; 241; 241
 CONFLICT; 242; 242
 CONFLICT; 288; 288
 SEQUENCE; 442 AA; 49116 MW; EA32PCFEF501331EE CRC64;

Query Match 91.9%; Score 57; DB 1; Length 442;
 Best Local Similarity 100.0%; Pred. No. 0.015; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

QY 1 RRPWCYGV 9
 |||||
 DB 131 RRPWCYGV 139

RESULT 4
 UROK_RAT

ID UROK_RAT STANDARD; PRT; 432 AA.
 AC P29598;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=Fischer 344;
 RX MEDLINE=92233409; Pubmed=1568219;
 RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
 RA Kellford R.F.;
 RT "Transcriptional and posttranscriptional activation of urokinase
 RT plasminogen activator gene expression in metastatic tumor cells";
 RL Cancer Res. 52:2489-2496(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Kidney;
 RA Rabbani S.A.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
 CC of two chains, A and B. The high molecular mass form contains a
 CC long chain A. Cleavage occurs after residue 156 in the low
 CC molecular mass form to yield a short A1 chain (by similarity).
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).

EMBL; X63434; CAA45028.1; --
 EMBL; X65651; CAA46601.1; --
 PIR; S24604; S18932.
 HSSP; P00749; 1KDU.
 MEROPS; S01.231; --
 InterPro; IPR009003; Cys Ser trypsin.
 InterPro; IPR006209; EGF-like.
 InterPro; IPR006210; IEGF.
 InterPro; IPR000001; Kringle.
 InterPro; IPR008293; Pept_SIA_uPA.
 InterPro; IPR01254; Peptidase_S1.
 InterPro; IPR01314; Peptidase_S1A.
 Pfam; PF00051; kringle; 1.
 Pfam; PF00089; trypsin; 1.
 PIRSF; PIRSF01144; UKR_plasm_act; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 ProDom; PD000395; Kringle; 1.
 SMART; SMO0130; KR; 1.
 SMART; SMO0020; Tryp_Spc; 1.
 PROSITE; PS00022; EGF_1; 1.
 PROSITE; PS01186; EGF_2; FALSE_NEG.
 PROSITE; PS50026; EGF_3; 1.
 PROSITE; PS00021; KRINGLE_1; 1.
 PROSITE; PS50070; KRINGLE_2; 1.
 PROSITE; PS50240; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
 PROSITE; PS00135; TRYPSIN_SER; 1.

KW Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
FT Kringling; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 20 177 CHAIN A (BY SIMILARITY).
FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
FT DOMAIN 27 63 EGF-LIKE.
FT DOMAIN 70 151 KRINGLE.
FT DOMAIN 152 178 CONNECTING PEPTIDE.
FT DOMAIN 179 432 SERINE PROTEASE.
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
FT DISULFID 210 226 BY SIMILARITY.
FT DISULFID 218 289 BY SIMILARITY.
FT DISULFID 314 383 BY SIMILARITY.
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 373 401 BY SIMILARITY.
FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
FT CONFLICT 16 16 N -> H (IN REF. 2).
FT CONFLICT 24 24 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C9 CRC64;

Query Match 90.3%; Score 56; DB 1; Length 432;
Best Local Similarity 88.9%; Pred. No. 0.021;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
Db 129 RRPWCYQI 137

RESULT 5
UROK_BOVIN STANDARD; PRT; 433 AA.
ID UROK_BOVIN STANDARD; PRT; 433 AA.
AC 005569; 028209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Arctic endothelium;
RX MEDLINE=93216119; PubMed=8385052;
RA Kaetzelmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schlemmer W.-D.,
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.,
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RL Inc. Daity J. 5:605-617(1995).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- INDUCTION: By retinoic acid.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.

CC -1- SIMILARITY: Contains 1 kringling domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; L03546; AAA51419.1; -;
DR EMBL; X85801; CAAS9796.1; -;
DR EMBL; J05560; J05560.
DR HSP; P00749; 1LMW.
DR MEROPS; S01.231; -;
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringling.
DR InterPro; IPR008293; Pept_S1A_uPA.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00051; kringling_1.
DR Pfam; PF00088; trypsin_1.
DR PIRSF; PIRSF00114; Urokinase_act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringling_1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS50020; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
FT Kringling; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 90.3%; Score 56; DB 1; Length 433;
Best Local Similarity 88.9%; Pred. No. 0.022;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRPWCYQV 9
Db 131 RRPWCYQI 139

RESULT 6
UROK_MOUSE

ID UROK MOUSE STANDARD; PRT; 433 AA.
AC P06869; 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RA MEDLINE=85179474; PubMed=2985383;
RA Bellin D., Vassalli J.-D., Comdepine C., Godeau F., Nagamine Y.,
RA Reich E., Koehler H.P., Duvoisin R.M.;
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
RT mouse urokinase-type plasminogen activator.";
RL Eur. J. Biochem. 148:225-232(1985).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=86163489; PubMed=2831940;
RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
RT "The murine urokinase-type plasminogen activator gene.";
RL Biochemistry 26:8270-8279(1987).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists
CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 156 in the low
CC molecular mass form to yield a short A1 chain (By similarity).
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02389; CAA26231.1; -;
DR EMBL; M17922; AAA40539.1; -;
DR PIR; A29420; UKMS.
DR HSSP; P00749; 1XDU.
DR MEROPS; S01.231; -;
DR MGD; MGI:97611; Plau.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR008293; Pept_S1A_UA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PIRSF; PIRSF001144; Utk_plasm act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.

DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 178 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 157 178 CHAIN A (BY SIMILARITY).
FT CHAIN 180 433 SHORT A CHAIN (A1).
FT CHAIN 28 64 CHAIN B (BY SIMILARITY).
FT DOMAIN 71 152 EGF-LIKE.
FT DOMAIN 153 179 KRINGLE.
FT DOMAIN 180 433 CONNECTING PEPTIDE.
FT DOMAIN 32 40 SERINE PROTEASE.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 54 63 BY SIMILARITY.
FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
SQ SEQUENCE 433 AA; 48268 MW; A99C35F625043F9 CRC64;
Query Match 85.5%; Score 53; DB 1; Length 433;
Best Local Similarity 77.8%; Pred. No. 0.069;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRPWCYQV 9
Db 130 KRPMCYQI 138

RESULT 7
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q9UKV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP SMPAIN-BALB/c;
RC Tsch H., Kataoka H., Koono H.;
RA "Mouse hepatocyte growth factor activator.";
RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2126753; PubMed=11032833;
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RA Yang J., Huan Y.;
RT "Activation of hepatocyte growth factor (HGF) by endogenous HGF
RT activator is required for metanephric kidney morphogenesis in
RT vitro.";
RL J. Biol. Chem. 276:15099-15106(2001).
CC -1- FUNCTION: Activates hepatocyte growth factor (HGF) by converting
CC it from a single chain to a heterodimeric form (By similarity).
CC -1- SUBUNIT: Dimer of a short chain and a long chain linked by a
CC disulfide bond (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted as an inactive single-chain
CC precursor and is then activated to a heterodimeric form (By
CC similarity).
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 2 EGF-like domains.

CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

CC -----
 CC EMBL: AF099017; AAF02469.1; -
 CC EMBL: AF224724; AAF34712.1; -
 CC HSSP: P00763; IDPO.
 CC MGD: MGI:1859281; Hgfac.
 CC DR INTERPRO: IPR003003; Cys_Ser_trypsin.
 CC DR INTERPRO: IPR000742; EGF_2.
 CC DR INTERPRO: IPR006209; EGF-like.
 CC DR INTERPRO: IPR000083; Fibronectin.
 CC DR INTERPRO: IPR000562; FN Type_II.
 CC DR INTERPRO: IPR006210; IEGF.
 CC DR INTERPRO: IPR000001; Kringle.
 CC DR INTERPRO: IPR001254; Peptidase_S1.
 CC DR INTERPRO: IPR001314; Peptidase_S1A.
 CC DR Pfam: PF00008; EGF_2.
 CC DR Pfam: PF00039; fn1; 1.
 CC DR Pfam: PF00040; fn2; 1.
 CC DR Pfam: PF00051; kringle; 1.
 CC DR Pfam: PF00089; trypsin; 1.
 CC DR PRINTS: PR00722; CHYMOTRYPSIN.
 CC DR PRINTS: PR00013; FNTYPEII.
 CC DR PRINTS: PR00018; KRINGLE.
 CC DR ProDom: PD000995; FN Type_II; 1.
 CC DR ProDom: PD000395; Kringle; 1.
 CC DR SMART: SM00181; EGF_2.
 CC DR SMART: SM00059; FN2; 1.
 CC DR SMART: SM00130; KR; 1.
 CC DR SMART: SM00020; TYP_SPC; 1.
 CC DR PROSITE: PS00022; EGF_1; 2.
 CC DR PROSITE: PS01186; EGF_2; 1.
 CC DR PROSITE: PS50026; EGF_3; 2.
 CC DR PROSITE: PS01253; FIBRONECTIN_1; 1.
 CC DR PROSITE: PS00023; FIBRONECTIN_2; 1.
 CC DR PROSITE: PS00021; KRINGLE_1; 1.
 CC DR PROSITE: PS50070; KRINGLE_2; 1.
 CC DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 CC DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 CC DR PROSITE: PS00135; TRYPSIN_SER; 1.
 CC KM Hydroxylase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 CC KM EGF-like domain; Repeat; Zymogen.
 CC FT SIGNAL 1 29
 CC FT PROPEL 30 369
 CC FT CHAIN 370 405
 CC FT
 CC FT CHAIN 406 653
 CC FT
 CC FT DOMAIN 105 145
 CC FT DOMAIN 157 195
 CC FT DOMAIN 197 237
 CC FT DOMAIN 238 276
 CC FT DOMAIN 283 364
 CC FT DOMAIN 406 653
 CC FT ACT_SITE 445 495
 CC FT ACT_SITE 495 495
 CC FT ACT_SITE 596 596
 CC FT DISULFID 105 130
 CC FT DISULFID 119 145
 CC FT DISULFID 161 172
 CC FT DISULFID 166 183
 CC FT DISULFID 185 194
 CC FT DISULFID 199 227

FT DISULFID 225 234 BY SIMILARITY.
 FT DISULFID 242 253 BY SIMILARITY.
 FT DISULFID 247 264 BY SIMILARITY.
 FT DISULFID 266 275 BY SIMILARITY.
 FT DISULFID 283 364 BY SIMILARITY.
 FT DISULFID 304 346 BY SIMILARITY.
 FT DISULFID 335 359 BY SIMILARITY.
 FT DISULFID 392 519 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 430 446 BY SIMILARITY.
 FT DISULFID 438 466 BY SIMILARITY.
 FT DISULFID 533 602 BY SIMILARITY.
 FT DISULFID 565 581 BY SIMILARITY.
 FT DISULFID 592 620 BY SIMILARITY.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 164 164 G -> W (IN REF. 2).
 SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;

Query Match 69.4%; Score 43; DB 1; Length 653;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPMCYV 7
 Db 343 RPMCYV 348

RESULT 8
 HGFA HUMAN STANDARD; PRT; 655 AA.
 ID HGFA HUMAN
 AC 004756; Q14726;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hepatocyte growth factor activator precursor (BC 3.4.21.-) (HGF
 DE activator) (HGFA).
 GN HGFA.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver, and Serum;
 RX MEDLINE=93252878; PubMed=768365;
 RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
 RA Kitamura N.;
 RT "Molecular cloning and sequence analysis of the cDNA for a human
 RT serine protease responsible for activation of hepatocyte growth
 RT factor. Structural similarity for the protease precursor to blood
 RT coagulation factor XII.";
 RT J. Biol. Chem. 268:10024-10028 (1993).
 RL [2]
 RN SEQUENCE OF 40-655 FROM N.A.
 RA Zhao S., Odeh C.;
 RC Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RL -1- FUNCTION: Activates hepatocyte growth factor (HGF) by
 CC converting it from a single chain to a heterodimeric form.
 CC -1- SUBUNIT: Dimer of a short chain and a long chain linked by a
 CC disulfide bond.
 CC -1- SUBCELLULAR LOCATION: Secreted as an inactive single-chain
 CC precursor and is then activated to a heterodimeric form.
 CC -1- TISSUE SPECIFICITY: Liver.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -1- CAUTION: It is uncertain whether Met-1 is the initiator.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M63990; AAA31595.1; -.
 CC PIR; J50600; JS0600.
 CC HSP; P98119; IAS1.
 CC MEROPS; S01.239; -.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR01254; Peptidase_S1.
 CC InterPro; IPR01314; Peptidase_S1A.
 CC Pfam; PF00051; Kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC PRODOM; PD000395; Kringle; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; TRYP_SPC; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS0070; KRINGLE_2; 1.
 CC PROSITE; PS50240; TRYPsin_DOM; 1.
 CC PROSITE; PS00134; TRYPsin_HIS; 1.
 CC PROSITE; PS00135; TRYPsin_SER; 1.
 CC K1M; Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
 CC Kringle; Signal; Multigene family.
 CC FT SIGNAL 1 36 POTENTIAL.
 CC FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
 CC FT DOMAIN 45 126 KRINGLE.
 CC FT ACT_SITE 142 394 SERINE PROTEASE.
 CC FT ACT_SITE 189 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT DISULFID 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT DISULFID 45 126 BY SIMILARITY.
 CC FT DISULFID 66 108 BY SIMILARITY.
 CC FT DISULFID 97 121 BY SIMILARITY.
 CC FT DISULFID 131 262 BY SIMILARITY.
 CC FT DISULFID 174 190 BY SIMILARITY.
 CC FT DISULFID 182 251 BY SIMILARITY.
 CC FT DISULFID 276 351 BY SIMILARITY.
 CC FT DISULFID 308 324 BY SIMILARITY.
 CC FT DISULFID 341 369 BY SIMILARITY.
 CC FT CARBOHYD 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
 CC SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52P3D81FCD CRC64;
 CC -----
 CC Query Match 64.5%; Score 40; DB 1; Length 394;
 CC Best Local Similarity 83.3%; Pred. No. 10;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 2 RPPCYV 7
 CC :|||||
 CC Db 105 KPMCYV 110
 CC -----
 CC RESULT 10
 CC URTB_DESCRO STANDARD; PRT; 431 AA.
 CC ID URTB_DESCRO
 CC AC P98121;
 CC DT 01-FEB-1996 (Rel. 33, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (OSPA
 CC beta).
 CC OS Desmodus rotundus (Vampire bat).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 CC OC Desmodontinae; Desmodus.
 CC OK NCBI_taxid=9430;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=2039036; PubMed=1937019;
 RA Kretschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schlenning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression";
 RL Gene 105:229-237(1991).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=9339059; PubMed=1309059;
 RA Schlenning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kretschmar J., Haendler B., Langer G., Baldus B., Witt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992)
 CC -1 FUNCTION: Probably essential to support the feeding habits of this
 CC exclusively haematophagous animal. Probable potent thrombolytic
 CC agent.
 CC -1 CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1 SUBUNIT: Monomer.
 CC -1 SIMILARITY: Belongs to peptidase family S1.
 CC -1 SIMILARITY: Contains 1 EGF-like domain.
 CC -1 SIMILARITY: Contains 1 Kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M63989; AAA31594.1; -.
 CC PIR; J50599; JS0599.
 CC HSP; P98119; IAS1.
 CC MEROPS; S01.239; -.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006209; EGF_Like.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR01254; Peptidase_S1.
 CC InterPro; IPR01314; Peptidase_S1A.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00051; Kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC PRODOM; PD000395; Kringle; 1.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; TRYP_SPC; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS00022; EGF_2; 1.
 CC PROSITE; PS0186; EGF_2; 1.
 CC PROSITE; PS50026; EGF_3; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS0070; KRINGLE_2; 1.
 CC PROSITE; PS50240; TRYPsin_DOM; 1.
 CC PROSITE; PS00134; TRYPsin_HIS; 1.
 CC PROSITE; PS00135; TRYPsin_SER; 1.
 CC K1M; Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
 CC Kringle; EGF-like domain; Signal; Multigene family.
 CC FT SIGNAL 1 36 POTENTIAL.
 CC FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
 CC FT DOMAIN 37 75 EGF-LIKE.
 CC FT DOMAIN 82 163 KRINGLE.
 CC FT DOMAIN 179 431 SERINE PROTEASE.
 CC FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC -----

```

FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 65 74 BY SIMILARITY.
FT DISULFID 82 163 BY SIMILARITY.
FT DISULFID 103 145 BY SIMILARITY.
FT DISULFID 134 158 BY SIMILARITY.
FT DISULFID 168 299 BY SIMILARITY.
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 288 BY SIMILARITY.
FT DISULFID 313 388 BY SIMILARITY.
FT DISULFID 345 361 BY SIMILARITY.
FT DISULFID 378 406 BY SIMILARITY.
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 431 AA; 48221 MW; 699B5E675B162CBF CRC64;

Query Match 64.5%; Score 40; DB 1; Length 431;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPMCIV 7
Db 142 KPMCTV 147

RESULT 11
UROK CHICK STANDARD; PRT; 434 AA.
ID UROK CHICK
AC P15120;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP MEDLINE=90110185; PubMed=2295632;
RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;
RT "The chicken urokinase-type plasminogen activator gene.";
RL J. Biol. Chem. 265:1339-1344(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL; J05187; AAA49131.1; -
CC EMBL; J05188; AAA49130.1; -
CC PIR; A35005; A35005.
CC HSSP; P00763; IDPO.
CC MEROPS; S01.231; -.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR008293; Pept_S1A_uPA.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF00089; trypsin; 1.

```

```

DR PIRSF; PIRSF001144; Urokinase_act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TYP_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS0186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 171 CHAIN A (BY SIMILARITY).
FT CHAIN 173 434 CHAIN B (BY SIMILARITY).
FT DOMAIN 36 72 EGF-LIKE.
FT DOMAIN 79 158 KRINGLE.
FT DOMAIN 159 172 CONNECTING PEPTIDE.
FT DOMAIN 173 434 SERINE PROTEASE.
FT DISULFID 40 48 BY SIMILARITY.
FT DISULFID 42 60 BY SIMILARITY.
FT DISULFID 62 71 INTERCHAIN (BY SIMILARITY).
FT DISULFID 162 296 BY SIMILARITY.
FT DISULFID 202 218 BY SIMILARITY.
FT DISULFID 210 285 BY SIMILARITY.
FT DISULFID 310 379 BY SIMILARITY.
FT DISULFID 342 358 BY SIMILARITY.
FT DISULFID 369 397 BY SIMILARITY.
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 434 AA; 49400 MW; BD881048D666A55 CRC64;

Query Match 64.5%; Score 40; DB 1; Length 434;
Best Local Similarity 71.4%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RPMCIV 8
Db 136 KPMCTV 142

RESULT 12
URTI DESRO STANDARD; PRT; 477 AA.
ID URTI DESRO
AC P98119;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA
DE alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Buthera; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kretschnar J., Haendler B., Langer G., Boidl W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]

```


RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schlemming W.-D., Alagon A., Boldon W., Bringmann P., Petri T.,
 RA Kaerzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RC TISSUE=Salivary gland;
 RX MEDLINE=9802274; PubMed=9354616;
 RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
 RA Schlemming W.D., Bode W.;
 RT "Catalytic domain structure of vampire bat plasminogen activator: a
 RT molecular paradigm for proteolysis without activation cleavage.";
 RL Biochemistry 36:13483-13493(1997).
 CC -1- FUNCTION: Probably essential to support the feeding habits of this
 CC exclusively haematophagous animal. Potent thrombolytic agent.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in
 CC plasminogen to form plasmin.
 CC -1- ENZYME REGULATION: Activity toward plasminogen is stimulated in
 CC the presence of fibrin I.
 CC -1- SUBUNIT: Monomer.
 CC -1- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,
 CC and the kringle domain apparently mediates fibrin-induced
 CC stimulation of activity.
 CC -1- SIMILARITY: Belongs to peptide family S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M63987; AAA31591.1; -;
 DR EMBL: M63986; AAA31592.1; -;
 DR PIR: J50597; J50597.
 DR PDB: 1A5I; 23-MAR-99.
 DR MEROPS: S01.232; -;
 DR GlycoSuiteDB: P98119; -;
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR000083; Fibrinctn1.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR003314; Peptidase_S1A.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00039; fn1_1.
 DR Pfam: PF00051; kringle_1.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00016; KRINGLE.
 DR ProDom: PD000395; Kringle_1.
 DR SMART: SM00181; EGF_1.
 DR SMART: SM00058; FN1_1.
 DR SMART: SM00130; KR_1.
 DR SMART: SM00020; TRYP_SPC_1.
 DR PROSITE: PS00023; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS50026; EGF_3; 1.
 DR PROSITE: PS01253; FIBRONECTIN_1; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS50070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPsin_DOM_1.
 DR PROSITE: PS00134; TRYPsin_HIS_1.
 DR PROSITE: PS00135; TRYPsin_SER_1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 36
 FT CHAIN 37 477
 FT DOMAIN 40 82
 FT DOMAIN 83 121
 FT DOMAIN 128 209
 FT DOMAIN 225 477
 FT ACT_SITE 272 321
 FT ACT_SITE 321 321
 FT ACT_SITE 428 428
 FT DISULFID 42 72
 FT DISULFID 70 79
 FT DISULFID 87 98
 FT DISULFID 92 109
 FT DISULFID 111 120
 FT DISULFID 128 209
 FT DISULFID 149 191
 FT DISULFID 180 204
 FT DISULFID 214 345
 FT DISULFID 257 273
 FT DISULFID 265 334
 FT DISULFID 359 434
 FT DISULFID 391 407
 FT DISULFID 424 452
 FT CARBOHYD 153 153
 FT CARBOHYD 398 398
 FT TRN 214 215
 FT STRAND 223 224
 FT TURN 226 227
 FT STRAND 230 231
 FT HELIX 234 236
 FT TURN 238 239
 FT STRAND 240 245
 FT STRAND 244 263
 FT TURN 264 265
 FT STRAND 266 269
 FT HELIX 271 273
 FT TURN 280 282
 FT STRAND 284 287
 FT TURN 297 298
 FT STRAND 300 309
 FT TURN 311 312
 FT TURN 315 317
 FT TURN 319 320
 FT STRAND 323 328
 FT STRAND 338 338
 FT TURN 339 340
 FT STRAND 341 341
 FT STRAND 345 345
 FT TURN 349 350
 FT TURN 355 356
 FT STRAND 358 363
 FT STRAND 366 366
 FT STRAND 374 374
 FT STRAND 379 385
 FT HELIX 388 390
 FT TURN 393 398
 FT TURN 403 404
 FT STRAND 405 409
 FT TURN 425 426
 FT TURN 428 429
 FT STRAND 431 436
 FT TURN 437 438
 FT STRAND 439 448
 FT TURN 455 456
 FT STRAND 459 463
 FT HELIX 464 467
 FT HELIX 468 474
 SQ SEQUENCE 477 AA; 53616 MW; AA06FD1739C10B5E CRC64;

N-LINKED (GLCNAC. . .)
 /FTid-CAR_000027.
 N-LINKED (GLCNAC. . .)
 /FTid-CAR_000028.

```

Query Match      64.5%; Score 40; DB 1; Length 477;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 RPMCYV 7
      :|||||
Db      188 KPMCYV 193

RESULT 13
UR12 DESRO      STANDARD; PRT; 477 AA.
AC p15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OC NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Krietzschmar J., Haendler B., Langer G., Boisdol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Salivary gland;
RX MEDLINE=90036867; PubMed=2509450;
RA Gadel S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
RT "Isolation, characterization, and cDNA cloning of a vampire bat
RT salivary plasminogen activator.";
RL J. Biol. Chem. 264:17947-17952(1989).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=9339059; PubMed=1309059;
RA Schleuning W.D., Alagon A., Boisdol W., Bringmann P., Petri T.,
RA Krietzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: Probably essential to support the feeding habits of this
CC exclusively haematophagous animal. Probable potent thrombolytic
CC agent.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- ENZYME REGULATION: Activity toward plasminogen is stimulated in
CC the presence of fibrin I.
CC -1- SUBUNIT: Monomer.
CC -1- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,
CC and the kringle domain apparently mediates fibrin-induced
CC stimulation of activity.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC EMBL; M63988; AAA31593.1; -
DR EMBL; J05082; AAA31596.1; -
DR PIR; A34363; A34369.
DR PIR; J05098; J05098.
DR HSSP; P98119; IASI.
DR MEROPS; S01.232; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle_S1.
DR InterPro; IPR001254; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50070; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KM Plasminogen activation: Hydrolyase; Serine protease; Glycoprotein;
KM kringle; EGF-like domain; signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 477
FT DOMAIN 40 82 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE.
FT DOMAIN 225 272 SERINE PROTEASE.
FT ACT_SITE 272 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 214 345
FT DISULFID 257 273
FT DISULFID 265 334
FT DISULFID 359 434
FT DISULFID 391 407
FT DISULFID 424 452
FT CARBOHYD 185 185
FT CARBOHYD 398 398
FT CARBOHYD 403 403
FT CONFLICT 417 417
FT CONFLICT 435 435
SQ SEQUENCE 477 AA; 53719 MW; 1748655C0DE5077C CRC64;
```

```

Query Match      64.5%; Score 40; DB 1; Length 477;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 RPMCYV 7
      :|||||
Db      188 KPMCYV 193
```



```

FT  CONFLICT 325 325 P -> A (IN REF. 1).
SQ  SEQUENCE 559 AA; 63122 MW; 8CCCE2BDB94514D9 CRC64;
Query Match 64.5%; Score 40; DB 1; Length 559;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 2 RPMCIV 7
Db 184 KPMCTV 189

RESULT 15
TPA_RAT STANDARD; PRT; 559 AA.
AC P19637;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor (BC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89170114; Pubmed=3148445;
RA My T., Leonardsson G., Hsueh A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator."
RL DNA 7:671-677 (1988).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90130448; Pubmed=2105315;
RA Feng P., Ohlsson M., Ny T.;
RT "The structure of the TATA-less rat tissue-type plasminogen activator
RT gene. Species-specific sequence divergences in the promoter predict
RT differences in regulation of gene expression."
RL J. Biol. Chem. 265:2022-2027 (1990).
CC -1- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
CC controlling plasmin-mediated proteolysis, it plays an important
CC role in tissue remodeling and degradation, in cell migration and
CC many other physiological events.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
CC bond.
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular.
CC -1- PM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: Binds to the kringle structure of the fibrin A
CC chain. Binding to fibrin enhances its catalytic activity.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31187; AAA42261.1; JOINED.
DR EMBL; M31188; AAA42261.1; JOINED.
DR EMBL; M31189; AAA42261.1; JOINED.
DR EMBL; M31190; AAA42261.1; JOINED.
DR EMBL; M31191; AAA42261.1; JOINED.
DR EMBL; M31192; AAA42261.1; JOINED.
DR EMBL; M31193; AAA42261.1; JOINED.
DR EMBL; M31194; AAA42261.1; JOINED.
DR EMBL; M31195; AAA42261.1; JOINED.
DR EMBL; M31196; AAA42261.1; JOINED.
DR EMBL; A19618; CAA01482.1; -.
DR PIR; A35029; A35029.
DR HSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1_1.
DR Pfam; PF00051; Kringle_2.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle_2.
DR SMART; SM00181; EGF_1.
DR SMART; SM00058; FN1_1.
DR SMART; SM00130; KR_2.
DR SMART; SM00020; TRYP_SPC_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01166; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM_1.
DR PROSITE; PS00134; TRYPSIN_HIS_1.
DR PROSITE; PS00135; TRYPSIN_SER_1.
DR KMW; Plasminogen activator; Hydrolyase; Serine protease; Glycoprotein;
DR KMW; Plasmin; Kringle; EGF-like domain; Repeat; Signal.
DR SIGNAL 1 17
DR PROPEP 18 29
DR CHAIN 30 559
DR CHAIN 30 306
DR CHAIN 309 559
DR DOMAIN 36 78
DR DOMAIN 79 117
DR DOMAIN 124 205
DR DOMAIN 213 294
DR DOMAIN 309 559
DR ACT_SITE 355 355
DR ACT_SITE 404 404
DR ACT_SITE 510 510
DR DISULFID 38 68
DR DISULFID 66 75
DR DISULFID 83 94
DR DISULFID 88 105
DR DISULFID 107 116
DR DISULFID 124 205
DR DISULFID 145 187
DR DISULFID 176 200
DR DISULFID 213 294
DR DISULFID 234 276
DR DISULFID 265 289
DR DISULFID 297 428
DR DISULFID 340 356
DR DISULFID 348 417
DR DISULFID 442 516

```

```

DR EMBL; M31187; AAA42261.1; JOINED.
DR EMBL; M31188; AAA42261.1; JOINED.
DR EMBL; M31189; AAA42261.1; JOINED.
DR EMBL; M31190; AAA42261.1; JOINED.
DR EMBL; M31191; AAA42261.1; JOINED.
DR EMBL; M31192; AAA42261.1; JOINED.
DR EMBL; M31193; AAA42261.1; JOINED.
DR EMBL; M31194; AAA42261.1; JOINED.
DR EMBL; M31195; AAA42261.1; JOINED.
DR EMBL; M31196; AAA42261.1; JOINED.
DR EMBL; A19618; CAA01482.1; -.
DR PIR; A35029; A35029.
DR HSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1_1.
DR Pfam; PF00051; Kringle_2.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle_2.
DR SMART; SM00181; EGF_1.
DR SMART; SM00058; FN1_1.
DR SMART; SM00130; KR_2.
DR SMART; SM00020; TRYP_SPC_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01166; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM_1.
DR PROSITE; PS00134; TRYPSIN_HIS_1.
DR PROSITE; PS00135; TRYPSIN_SER_1.
DR KMW; Plasminogen activator; Hydrolyase; Serine protease; Glycoprotein;
DR KMW; Plasmin; Kringle; EGF-like domain; Repeat; Signal.
DR SIGNAL 1 17
DR PROPEP 18 29
DR CHAIN 30 559
DR CHAIN 30 306
DR CHAIN 309 559
DR DOMAIN 36 78
DR DOMAIN 79 117
DR DOMAIN 124 205
DR DOMAIN 213 294
DR DOMAIN 309 559
DR ACT_SITE 355 355
DR ACT_SITE 404 404
DR ACT_SITE 510 510
DR DISULFID 38 68
DR DISULFID 66 75
DR DISULFID 83 94
DR DISULFID 88 105
DR DISULFID 107 116
DR DISULFID 124 205
DR DISULFID 145 187
DR DISULFID 176 200
DR DISULFID 213 294
DR DISULFID 234 276
DR DISULFID 265 289
DR DISULFID 297 428
DR DISULFID 340 356
DR DISULFID 348 417
DR DISULFID 442 516

```

FT DISULFID 474 490 BY SIMILARITY.
 FT DISULFID 506 534 BY SIMILARITY.
 FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 380 380 E -> K (IN REF. 1).
 SQ SEQUENCE 559 AA; 62903 MM; 7DBD3809CIDIC921 CRC64;

Query Match 64.5%; Score 40; DB 1; Length 559;
 Best Local Similarity 83.3%; Pred. No. 14;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPNQYV 7
 :|||
 Db 184 KPNQYV 189

Search completed: March 18, 2004, 11:17:01
 Job time : 0.612691 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 11:09:25 ; Search time 2.73533 Seconds
(without alignments)
1153.535 Million cell updates/sec

Title: US-10-076-421-5
Perfect score: 62
Sequence: 1 RRPWCYVQVQ 10

Scoring table: BLAST62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	91.9	154	4	Q96SE8
2	57	91.9	214	6	Q9XT70
3	57	91.9	433	6	Q8MTL0
4	57	91.9	433	6	Q8MHY7
5	56	90.3	157	6	Q9TVAS
6	53	85.5	231	11	Q8C6L2
7	46	74.2	1029	5	Q81252
8	43	69.4	608	16	Q9KOP9
9	43	69.4	653	11	Q8VCS4
10	43	69.4	852	5	Q8IKW0
11	41	66.1	97	10	P82624
12	41	66.1	201	16	Q7U319
13	41	66.1	234	4	Q86WZ2
14	41	66.1	263	4	Q00318
15	41	66.1	263	4	Q8NCJ9
16	41	66.1	263	4	Q96FE7

17	41	66.1	517	11	Q8K0D2	Q8K0D2 mus musculus
18	40	64.5	202	13	Q90675	Q90675 gallus galli
19	40	64.5	213	11	Q811Z3	Q811Z3 mus musculus
20	40	64.5	263	13	Q7SXB3	Q7SXB3 brachydactyl
21	40	64.5	264	11	Q811Z2	Q811Z2 mus musculus
22	40	64.5	264	11	Q7TMJ8	Q7TMJ8 mus musculus
23	40	64.5	272	11	Q8BWS2	Q8BWS2 mus musculus
24	40	64.5	291	4	Q7Z7N2	Q7Z7N2 homo sapien
25	40	64.5	304	16	Q9A3Q4	Q9A3Q4 caulobacter
26	40	64.5	411	11	Q8R0L1	Q8R0L1 mus musculus
27	40	64.5	516	4	Q9BU99	Q9BU99 homo sapien
28	40	64.5	519	10	Q8IACS	Q8IACS arabidopsis
29	40	64.5	519	10	Q84W90	Q84W90 arabidopsis
30	40	64.5	560	4	Q14520	Q14520 homo sapien
31	40	64.5	562	4	Q86YK8	Q86YK8 homo sapien
32	40	64.5	564	6	Q8MKB1	Q8MKB1 oryctolagus
33	40	64.5	566	10	Q911L7	Q911L7 arabidopsis
34	40	64.5	567	4	Q13208	Q13208 homo sapien
35	40	64.5	947	13	Q8AXY6	Q8AXY6 gallus galli
36	40	64.5	1109	5	Q869M0	Q869M0 dictyostell
37	39	62.9	60	4	Q9UKJ7	Q9UKJ7 homo sapien
38	39	62.9	113	4	Q9URJ7	Q9URJ7 homo sapien
39	39	62.9	215	13	Q42341	Q42341 gallus galli
40	39	62.9	285	4	Q8TCE2	Q8TCE2 homo sapien
41	39	62.9	280	4	Q02935	Q02935 homo sapien
42	39	62.9	296	4	Q14519	Q14519 homo sapien
43	39	62.9	374	17	Q8Z244	Q8Z244 pyrobaculum
44	39	62.9	421	13	Q8AXX3	Q8AXX3 xenopus lae
45	39	62.9	454	6	Q46506	Q46506 papio hamad

ALIGNMENTS

RESULT 1
ID Q96SE8 PRELIMINARY; PRT; 154 AA.
AC Q96SE8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029537; AAK38734.1; -;
DR GO; GO:0016301; F-kinase activity; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PSS0070; KRINGLE_2; 1.
 KM Glycopolypeptide; Kinase; Kringle.
 SQ SEQUENCE 154 AA; 17305 MW; AACCF2FCFF505572 CRC64;

Query Match 91.9%; Score 57; DB 4; Length 154;
 Best Local Similarity 100.0%; Pred. No. 0.017;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPWCYQV 9
 |||||
 DB 129 RRPWCYQV 137

RESULT 2

09XT70 PRELIMINARY; PRT; 214 AA.

AC 09XT70 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 OS Urokinase-type plasminogen activator (Fragment).
 DE Urokinase-type plasminogen activator (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA yin J., Idel S.;
 RT "Partial mRNA of rabbit uPA."
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL: AF097647; AAD39351.1; -
 DR HSRP; P00749; IETN.
 DR MEROPS; S01.231; -
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_Trypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRODOM; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRY_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_2; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 KM Glycopolypeptide; Kinase; Kringle; Protease; Serine protease.
 FT NON TER 214
 SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32BD7E CRC64;

Query Match 91.9%; Score 57; DB 6; Length 214;
 Best Local Similarity 100.0%; Pred. No. 0.024;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPWCYQV 9
 |||||
 DB 43 RRPWCYQV 51

RESULT 3

08MIL0 PRELIMINARY; PRT; 433 AA.

AC 08MIL0 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Urokinase-type plasminogen activator.
 GN PLAU.

OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=2215945; Pubmed=12149463;
 RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
 RA Dicker D.A.;
 RT "Increased expression of urokinase during atherosclerotic lesion
 development causes arterial constriction and lumen loss, and
 accelerates lesion growth."
 RT Proc Natl Acad Sci U S A. 99:10665-10670(2002).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; AY12285; AAM83187.1; -
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_Trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR008295; Pept_S1A_uPA.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRY_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR PIRSF; PIRSF001144; Utk_plasm_act; 1.
 KM Glycopolypeptide; Kinase; Kringle; Protease; Serine protease.
 SQ SEQUENCE 433 AA; 48375 MW; 65E64F364155490 CRC64;

Query Match 91.9%; Score 57; DB 6; Length 433;
 Best Local Similarity 100.0%; Pred. No. 0.046;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRPWCYQV 9
 |||||
 DB 131 RRPWCYQV 139

RESULT 4

08MHY7 PRELIMINARY; PRT; 433 AA.

AC 08MHY7 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Urokinase-type plasminogen activator.

GN UROKINASE.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Sugiki M., Yoshida E., Anai K., Maruyama M.;

Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA Yano W., Watanabe M.,
RT "oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL: AY029517; AAK40239.1; -
DR EMBL: AB087224; BAC02685.1; -
DR GO: GO:0004263; F:chymotrypsin activity; IEA.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004295; F:trypsin activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys Ser trypsin.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR01254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR008293; Pept_S1A_upa.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYF_SPC; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR PIRSF: PIRSF001144; Utk_plasm_act; 1.
DR GlycoProtein: Hydroxylase; Kinase; Kringle; Protease; Serine protease.
KW SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EB CRC64;
Query Match 91.9%; Score 57; DB 6; Length 433;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRPWCYQV 9
Db 131 RRPWCYQV 139
RESULT 5
Q9TVAB PRELIMINARY; PRT; 157 AA.
AC Q9TVAB;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RA MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Querengesser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle."
RL J. Anim. Sci. 79:94-107(2001)
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL: AF144761; AAD30301.1; -
DR HSSP: P00749; IURK.
DR GO: GO:0016301; F:kinase activity; IEA.

InterPro: IPR006209; EGF_like.
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; Kringle; 1.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
KW GlycoProtein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A76BD6C72C1FBFB7 CRC64;
Query Match 90.3%; Score 56; DB 6; Length 157;
Best Local Similarity 88.9%; Pred. No. 0.026;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRPWCYQV 9
Db 97 RRPWCYQV 105
RESULT 6
Q8C6L2 PRELIMINARY; PRT; 231 AA.
ID Q8C6L2;
AC Q8C6L2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK054349; BAC35743.1; -
DR PIR: PT0534; PT0534.
DR GO: GO:0004295; F:trypsin activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys_Ser trypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Peptidase_S1.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
FT NON_TER 231
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25B8980A682737E2 CRC64;
Query Match 85.5%; Score 53; DB 11; Length 231;
Best Local Similarity 77.8%; Pred. No. 0.13;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRPWCYQV 9
Db 130 RRPWCYQV 138
RESULT 7

```

Q81252
ID Q81252 PRELIMINARY; PRT; 1029 AA.
AC Q81252;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein with putative LCCL domain.
GN PFA0445W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=36329;
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D., Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooke K., Buckee C.O., Burrows C., Cherevach I., Chillingworth C., Chillingworth T., Christodoulou Z., Clark L., Clark R., Cotton C., Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J., Feltwell T., Goble A., Goodhead I., Gilliam R., Hamlin N., Hance Z., Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P., Humphrey S., Jagels K., James K.D., Johnson D., Kerhornou A., Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N., Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L., Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E., Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M., Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K., Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J., Sultson J.E., Craig A., Newbold C., Barrell B.G.
RA "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RT Nature 419:527-531 (2002).
RL EMBL; AL031745; CAD49049.1; -.
DR InterPro; IPR004043; LCCL_dom.
DR Pfam; PF03815; LCCL; 1.
DR SMART; SM00603; LCCL; 1.
DR PROSITE; PS50820; LCCL; 1.
KW Hypothetical protein.
SQ SEQUENCE 1029 AA; 118532 MW; 6512F6CE1B6D3725 CRC64;

Query Match 74.2% Score 46; DB 5; Length 1029;
Best Local Similarity 77.8%; Pred. No. 8.9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RPMCVCVQVQ 10
Db 73 REMCVCEVQ 81

RESULT 8
Q9KOP9 PRELIMINARY; PRT; 608 AA.
AC Q9KOP9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PvCA protein.
GN VCI1949.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrrio.
NCBI_TaxID=666;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., Ermolaeva M.D., Vamathevan J., Baas S., Qin H., Dragoi I., Sellers P., McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;
RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio

```

```

RT cholerae."
RL Nature 406:477-483 (2000).
DR EMBL; AE004270; AAF95097.1; -.
DR PIR; G82137; G82137.
DR TIGR; VCI1949; -.
DR InterPro; IPR007817; DIT1_PvCA.
DR Pfam; PF05141; DIT1_PvCA; 1.
KW Complete proteome.
SQ SEQUENCE 608 AA; 70459 MW; BAF8EEF3F11C4299 CRC64;

Query Match 69.4% Score 43; DB 16; Length 608;
Best Local Similarity 62.5%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPMCVCVQV 9
Db 306 KPMCVLQDL 313

RESULT 9
Q8VCS4 PRELIMINARY; PRT; 653 AA.
AC Q8VCS4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RX NBI;
RA "Sequence from N.A.
RC Tissue=Liver;
RA Strausberg R.;
Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSP; P00761; IAN1.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR000562; FN_Type_1I.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001314; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00054; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNYPEIT.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000995; FN_Type_1I; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM000130; KR; 1.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00023; EGF_1; 2.
DR PROSITE; PS01166; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.

```

DR PROSITE; PSS0070; KRINGLE 2; 1.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PSS0134; TRYPSIN_HIS; 1.
 DR PROSITE; PSS0135; TRYPSIN_SER; 1.
 KM Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
 KM Kringle; Protease; Serine protease.
 KM SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;
 Query Match
 Best Local Similarity 69.4%; Score 43; DB 11; Length 653;
 Pred. Local 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RPPCYVQV 7
 Db 343 RPPCYV 348
 RESULT 10
 Q8IKW0 PRELIMINARY; PRT; 852 AA.
 AC Q8IKW0;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Hypothetical protein.
 GN PF14_0491.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RX MEDLINE=2255705; PubMed=1236864;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyte S.,
 RA Chan W.-S., Nene V., Shallow S.D., Suh B., Peterson J., Angiuoli S.,
 RA Pereira M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
 RA Martin D.M.A., Fairhead A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrett B.;
 RT "genome sequence of the human malaria parasite Plasmodium
 falciparum."
 RL Nature 419:498-511(2002).
 DR EMBL: AB014824; AAA37104.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 852 AA; 98539 MW; 77FD6608825C113C CRC64;
 Query Match
 Best Local Similarity 69.4%; Score 43; DB 5; Length 852;
 Pred. Local 100.0%; Pred. No. 25;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 RPPCYVQVQ 10
 Db 60 KEMCYVEVQ 68
 RESULT 11
 P82624 PRELIMINARY; PRT; 97 AA.
 AC P82624;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE SCRL5 precursor.
 GN SCRL5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Becker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hutzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kvan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Maltfecher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
 RA Utechtack T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RL Nature 408:816-820(2000).
 RN [2]
 RP IDENTIFICATION, AND TISSUE SPECIFICITY.
 RX PubMed=11437247;
 RA Vanoosthuyse V., Mieg C., Dumas C., Cock J.M.;
 RT "Two large Arabidopsis thaliana gene families are homologous to the
 RT Brassica gene superfamily that encodes pollen coat proteins and the
 RT male component of the self-incompatibility response."
 RL Plant Mol. Biol. 46:17-34(2001).
 CC -1- TISSUE SPECIFICITY: FLOWER BUDS AND STEMS.
 DR EMBL: AC018908; -; NOT_ANNOTATED_CDS.
 KW SIGNAL.
 FT SIGNAL 1 22
 FT CHAIN 23 97 SCRL5.
 SQ SEQUENCE 97 AA; 11119 MW; 66CA0B925B7D8F54 CRC64;
 Query Match
 Best Local Similarity 66.1%; Score 41; DB 10; Length 97;
 Pred. Local 100.0%; Pred. No. 72;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RPPCYVQVQ 10
 Db 35 RPPCPCKIQ 44
 RESULT 12
 Q7U319 PRELIMINARY; PRT; 201 AA.
 AC Q7U319;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DE Urease accessory protein G.
 GN UREG OR SYN2443.
 OS Synechococcus sp. (strain WH8102).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxID=84588;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22825697; PubMed=12917641;
 RA Palenik B., Brahamsa B., Larimer F.W., Land M., Hauser L., Chain P.,
 RA Lamerdin J., Regala W., Allen E.E., McCarron J., Paulsen I.,
 RA Dufresne A., Paternsky F., Webb E.A., Waterbury J.;
 RT "The genome of a motile marine Synechococcus."
 RL Nature 424:1037-1042(2003).
 DR EMBL: BK569695; CAB08958.1; -
 KW Complete proteome.
 SQ SEQUENCE 201 AA; 21557 MW; 7096BF2C566CF597 CRC64;
 Query Match
 Best Local Similarity 66.1%; Score 41; DB 16; Length 201;
 Pred. Local 100.0%; Pred. No. 14;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RPPWCYOVQ 10
 ||||: :
 Db 174 RPPWCYV 182

RESULT 13

Q86YV2 PRELIMINARY; PRT; 234 AA.
 AC O86YV2;
 DT 01-JUN-2003 (TREMblrel. 24, Created)
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE HGFL(8) protein.
 GN HGFL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chiang H., Chang M.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF528079; AA033762.1; -.
 DR InterPro; IPR000001; Kringie.
 DR Pfam; PF00051; Kringie; 1.
 DR PRINTS; PR00018; KRINGIE.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGIE_1; 1.
 DR PROSITE; PS50070; KRINGIE_2; 1.
 SQ SEQUENCE 234 AA; 25320 MW; C78F64DBD18DCDC CRC64;

Query March 66.1%; Score 41; DB 4; Length 234;
 Best Local Similarity 85.7%; Pred. No. 17; Mismatches 1; Indels 0; Gaps 0;
 Matches 6; Conservative 0;

Qy 1 RPPWCYV 7
 |||||
 Db 78 RPPWCYV 84

RESULT 14
 O00318 PRELIMINARY; PRT; 263 AA.
 AC O00318;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE WDSCC; DJ515N1.2 protein.
 GN WDSCC; DJ515N1.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Du Z., Scheet P., Harper M.;
 RT "The sequence of H. sapiens PAC clone RP3-515N1."
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 KRINGIE DOMAIN.
 DR EMBL; AC002073; AAB54054.1; -.
 DR HSSP; P00749; IKDU.
 DR InterPro; IPR000001; Kringie.
 DR Pfam; PF00051; Kringie; 1.
 DR PRINTS; PR00018; KRINGIE.
 DR ProDom; PD000395; Kringie; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGIE_1; 1.
 DR PROSITE; PS50070; KRINGIE_2; 1.

KW Glycoprotein; Kringie.
 SQ SEQUENCE 263 AA; 28248 MW; 197C3EBE8E54A242 CRC64;

Query March 66.1%; Score 41; DB 4; Length 263;
 Best Local Similarity 85.7%; Pred. No. 18; Mismatches 1; Indels 0; Gaps 0;
 Matches 6; Conservative 0;

Qy 1 RPPWCYV 7
 |||||
 Db 78 RPPWCYV 84

RESULT 15

Q86NCJ9 PRELIMINARY; PRT; 263 AA.
 AC Q86NCJ9;
 DT 01-OCT-2002 (TREMblrel. 22, Created)
 DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Hypothetical protein FLJ30207.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue-Mammary gland;
 RA Isogai T., Ota T., Nishikawa T., Hayaishi K., Otsuki T., Sugiyama T.,
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
 RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
 RT "MDO human cDNA sequencing project."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 KRINGIE DOMAIN.
 DR EMBL; AK074688; BAC11140.1; -.
 DR InterPro; IPR000001; Kringie.
 DR Pfam; PF00051; Kringie; 1.
 DR PRINTS; PR00018; KRINGIE.
 DR ProDom; PD000395; Kringie; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGIE_1; 1.
 DR PROSITE; PS50070; KRINGIE_2; 1.
 KW Hypothetical protein; Glycoprotein; Kringie.
 SQ SEQUENCE 263 AA; 28104 MW; 73A9294D5142C6C3 CRC64;

Query March 66.1%; Score 41; DB 4; Length 263;
 Best Local Similarity 85.7%; Pred. No. 18; Mismatches 1; Indels 0; Gaps 0;
 Matches 6; Conservative 0;

Qy 1 RPPWCYV 7
 |||||
 Db 78 RPPWCYV 84

Search completed: March 18, 2004, 11:19:22
 Job time : 3.73523 secs